Sequence specific DNA recognition by (cytosine-5)-methyltransferases

Ph.D Thesis

Tamás Raskó

Supervisor: Dr Antal Kiss

2003

Institute of Biochemistry
Biological Research Centre
Hungarian Academic of Science

Introduction

The Ph.D. thesis consists of two different projects. In the first section, I summarized results of a study on DNA bending induced by five DNA (cytosine-5)-methyltransferases. The second part of the thesis, deals with construction of *Sss*I methyltransferase variants, suitable for coupling with oligodeoxynucleotid (ODN) and peptid nucleic acid (PNA).

DNA (cytosine-5)-methyltransferases (C5-MTase) catalyze the transfer of a methyl group from *S*-adenosyl-L-methionine (AdoMet) to the C5 carbon of cytosine in specific sequences. These enzymes play an important role in several biological processes e.g. restriction-modification in bacteria, differentiation, regulation of gene expression in eukaryotes and carcinogenesis.

C5-MTases act as monomers. Our knowledge about C5-MTases is, to large extent, based on studies with bacterial enzymes. Eukaryotic C5-MTases are larger proteins but the sequence homology they share with bacterial C5-MTases and the available experimental data suggest that they act using the same catalytic mechanism. Bacterial C5-MTases share a common architecture, they contain ten conserved sequence motifs and a so-called variable region, which is responsible for sequence-specific DNA recognition.

Bending of the DNA duplex has proven a significant aspect of its interaction with many proteins. Several DNA binding proteins utilize bending as part of the recognition mechanism. Footprinting studies with the M.*Bsp*RI (cytosine-5)-methyltransferase suggested that the enzyme might induce a distortion, perhaps a bend in substrate DNA. To test this hypothesis, we performed mobility shift assays with circularly permutated DNA fragments. After the first surprising results with M.*Bsp*RI and M.*Hae*III, we extended the studies to three other C5-MTases.

In 2001, our group, as a member of an international consortium, won a research grant of the European Fifth Framework Programme for development of methyltransferases with programmable specifities. The ability to methylate DNA sequences playing a role in gene regulation may help to understand information generated by genomic sequencing. This task may be achievable by generation of methyltransferases with "programmable" specificities. Oligodeoxynucleotides (ODN) and peptid nucleic acid oligomers (PNA) have the capacity to form complexes with DNA. It is assumed that if ODN or PNA is covalently bound to a methyltransferase, the specifity of the conjugate will be partly determined by the sequence of the PNA/ODN part. The aim is to introduce PNA/ODN-methylase conjugates into cultured cells and disrupt transcription of certain genes by selective methylation of their promoter. The PNA/ODN's of the conjugate will make specifity of the methyltransferase programmable.

For coupling of PNA/ODN to a methyltransferase, the structure of the enzyme must be modified depending on the specific requirements of the coupling reactions. Coupling using the bifunctional crosslinker requires a Cys residue placed at or close to the N or the C terminus of the molecule. Because any thiol group in the protein may react with the reagent, the protein should ideally not contain any other cysteine. Native chemical peptide ligation is dependent on a free cysteine thiol group at the amino terminus of the target protein, internal Cys residues, however, do not react. The *Spiroplasma* C5-MTase, M.*Sss*I was choosen as model enzyme for coupling with ODN and PNA. M.*Sss*I has the lowest specificity (CG) among procaryotic C5-MTases, thus sequence specificity of the envisaged conjugates will be largely dependent on the “programming” ODNs or PNAs.

Methods

- Plasmid DNA purification
- DNA fragment analysis by agarose and poliacrylamide gelelectrophoresis
- Protein analysis by SDS-poliacrylamide gelelectrophoresis
- DNA purification from agarose and poliacrylamide-gel
- Introduction of plasmid DNA into *Escherichia coli* cells by transformation
- Use of phosphorylated and non-phosphorylated linkers
- PCR reactions and T/A cloning
- Site-directed mutagenesis
- Overexpression and purification of affinity-tagged M.*Sss*I variants
- Preparation of crude cell extract
- Measurement of methyltransferase activity

Results and discussion

DNA bending induced by (cytosine-5)- methyltransferases

1, For preparation of circularly permutated set of fragments, four plasmids (pBend-GGCC, pBend-GGWCC, pBend-GCGC, pBend-CCGG) were constructed. ^{32}P-labeled DNA fragments were synthesized *in vitro* by PCR amplification of the polylinker segment of pBend plasmids. The permutated set of fragments was generated by cleaving the PCR product with restriction enzymes having cleavage sites in the tandem repeats.

2, We performed mobility shift assays with five C5-MTases using a circularly permutated sets of DNA fragments. The calculated bend values were: M.*Bsp*RI (enzime specificity GG<u>C</u>C, the methylated cytosine is underlined) -

46-50°, M.*Hae*III (GG<u>C</u>C) - 40-43°, M.*Sin*I (GG$^{A}/_{T}$<u>C</u>C) - 33-36°, M.*Sau*96I (GGN<u>C</u>C) - 52-57°, M.*Hpa*II (C<u>C</u>GG) - 30°, M.*Hha*I (G<u>C</u>GC) 13°.

3, To obtain additional evidence for protein-induced DNA curving, we performed phasing experiments to identify direction of the DNA bend induced by *Bsp*RI and *Hae*III methyltransferases. This method relates complex mobility to the helical phasing between the protein-induced bend and a sequence-directed intrinsic bend on the same DNA fragment. Autoradiography of the gels with complexes formed between the methyltransferases and the labeled fragments showed that the relative mobility was a function of the distance between the $(A/T)_5$-tracts and the GGCC site. The pattern of the relation between mobility and the distance of the bend centers suggests that the direction of the bend induced by *Bsp*RI/*Hae*III methyltransferases-induced bend is toward the minor groove.

Conclusions

Our results and data from other laboratories show a correlation between the bending properties and the recognition specificities of (cytosine-5)-methyltransferases: enzymes recognizing a cytosine 3' to the target cytosine tend to induce greater bends than enzymes with guanine in this position.

What function DNA bending could have in the action of a DNA methyltransferase? In addition to possible a role in sequence-specific DNA recognition, it may play a role in base flipping. During base flipping the target base is rotated out from the helix by 180°. A connection between bending and base pair opening was indicated by molecular modeling. It was shown that DNA bending facilitates base pair opening and that, conversely, once a base pair is disrupted, DNA can bend very easily. This leads to the question whether

bending is a cause or a consequence of base flipping. One of the alternatives is that DNA bending is a mechanism by which the enzyme overcomes the forces holding the target base in the helix. Alternatively, bending may follow base flipping in time and its role may be to stabilize the DNA structure after the base has flipped out. The second model seems to be more likely. Firstly, base stacking energy values published in the literature do not support the first model. Using energy levels derived from quantum-chemical calculations for B DNA, the total (intrastrand plus interstrand) stacking energies between a cytosine base and its neighboring base pairs in an NCC context are –6.82, -5.52, -7.81, and -5.75 kcal/mol for sequences with N=A,C,G and T bases, respectively. The stacking energy for cytosine in a CCG context is –6.17 kcal/mol. These values do not explain for example, why M.*Msp*I induces a much greater bend than M.*Hpa*II. Secondly, structural models suggest that C5-MTases flip the cytosine through the minor groove. Molecular modeling indicates that base rotation toward the minor groove is energetically much less favorable than toward the major groove, because it leads to steric crowding of the atoms involved in the base pair hydrogen bonding. Therefore, if C5-MTase-induced bending is toward the minor groove, as our phasing experiments with M.*Hae*III seem to indicate, bending should inhibit rather than facilitate base rotation. We suggest that bending follows base flipping and its role would be to stabilize the DNA helix in the recognition complex. A recent molecular modeling study based on free-energy calculations indicates that base rotation occurs *via* the major groove. Flipping through the major groove would be compatible with a bending-first-model, in which bending towards the minor groove facilitates base pair opening and flipping.

Construction of M.SssI variants for coupling with PNA or ODN

Methyltransferase variants without affinity-tag

First, we attempted to modify the original pCAL7 plasmid (obtained from New England Biolabs), which encodes wild type M.*Sss*I. We constructed three mutant variants:

Gly386Cys mutant

Cys368Ala mutant

Cys368Ala + Gly386Cys double mutant

Unfortunatelly, neither the wild type, nor the mutant enzymes could be overproduced and purified to homogenity. To facilitate purification, we decided to construct variants, in which M.*Sss*I is fused to an affinity tag.

Affinity-tagged variants

First we tried to use the expression plasmid vectors pFLAG-MAC and pET3His to construct plasmids that overexpress M.*Sss*I. These approaches were not successful, because we did not get stabile clones.

Stabile overexpression clones were obtained only when the plasmid vector pBAD24 was used. All plasmids listed below are based on pBAD24.

For purposes of the coupling reaction using native chemical peptid ligation we have constructed two M.*Sss*I variants carrying the FLAG octapeptide at the N-terminus (FLAG-M.SssI and FLAG-M.Sss-74), and purified them to homogenity using a FLAG affinity column. The FLAG-M.SssI-74 variant differs from FLAG-M.SssI by the C368A mutation and the C-terminal Cys. The FLAG tag could be removed by enterokinase cleavage resulting in N-terminal cysteine. Because of its C-terminal Cys, this variant

could be, in principle, suitable for coupling using the bifunctional reagent without prior removal of the N-terminal affinity tail.

For the purpose of more effective purification we generated a "combined variant" (FLAG-M.Sss-6xHis-Cys), which carried the FLAG octapeptide at the N-terminus and a 6xHis-Cys peptide at the C-terminus. Advantage of this double-tagged variant is, that the second affinity tag will allow separation of the conjugate from free ODN/PNA.

In another approach, we mutated Ser2 to cysteine and cloned the modified gene in pBAD24 vector. This construct was made in the hope that the N-terminal formil-methionine will be removed enzymatically in *E.coli* cells resulting in N-terminal cysteine. To facilitate purification of this variant, an oligonucleotide encoding a 6xHis-peptide tag was added to the C-terminus of the *Sss*I methyltransferase gene.

During our work with M.*Sss*I clones we often observed that the cells died after induction of protein production. We thought that the reason of toxicity was extensive methylation of host DNA. To test this assumption, we generated an inactive variant of M.*Sss*I by mutating the codon of the catalytic cystein (C141S). Suprisingly, this variant was also toxic for cells after protein induction. We suggest that the explanation of toxicity is the binding of the *Sss*I MTase to CG sites which then blocks replication and transcription.

In the course of this work, several plasmid expression vectors were tried to overexpress M.*Sss*I and its mutant derivates. The only vector to give substantial overproduction was pBAD24. We think that the key factor for success with pBAD24 could be tight regulation of *araBAD* promoter.

One of the aims of our work is to purify M.*Sss*I for X-ray crystallographic study. We hope that studies of M.*Sss*I structure will help to understand the catalytic mechnism of the human methyltransferase which has similar specificity.

Publications and presentations

publications:

Kiss, A. , Pósfai, Gy., Zsurka, G., **Raskó, T**. and Venetianer, P. 2001. Role of DNA minor groove interactions in substrate recognition by the M.*Sin*I and M.*Eco*RII DNA (cytosine-5) methyltransferases. *Nucleic Acids Res*. 29: 3188-3194.

Simoncsits A., Tjornhammar, M.L., **Raskó, T**., Kiss, A.and Pongor S. 2001. Covalent joining of the subunits of a homodimeric type II restriction endonuclease: single-chain *Pvu*II endonuclease. *J. Mol. Biol*. 309: 89-97.

Raskó, T., Finta, C.and Kiss, A. DNA bending induced by DNA (cytosine-5) methyltransferases. 2000. *Nucleic Acids Res*. 28: 3083-91.

presentations:

Kiss, A., Finta, Cs. és **Raskó, T**. 1998, Budapest. Analysis of the *Bsp*RI methyltransferase-DNA interaction. Howard Huges Medical Institute. Meeting of International Research Scholars.

Raskó, T., Finta, Cs. és Kiss, A. 1999, Straub Days, Szeged. DNA induced by C5-cytosine methyltransferases.

Raskó, T. Finta, Cs. és Kiss, A. 2000, Sopron. DNA induced by C5-cytosine methyltransferases.

Kiss, A., Zsurka, G., Pósfai, Gy., **Raskó, T**., Venetianer, P. 2001, Sárospatak. Minor groove interactions in substrate recognition by C5-cytosine methyltransferases.

Tamás Raskó, Csaba Finta and Antal Kiss. 2001, Szeged. DNA bending induced by DNA (cytosine-5)-methyltransferases. Hungarian-Polish Ph.D conference.

Tamás Raskó, Slaska-Kiss Krystyna és Antal Kiss. 2001, Kréta. 2002, Newcastle. 2003, Verbania. 2003, Berlin. Engineering M.*Sss*I variants for terminal coupling to oligonucleotides or PNA. Meetings of the FP5 Programme "QUALITY OF LIFE AND MANAGEMENT OF LIVING RESOURCES".