Flexibility of the Linker between the Domains of DNA Methyltransferase SsoII Revealed by Small-Angle X-Ray Scattering: Implications for Transcription Regulation in SsoII Restriction–Modification System

Petr V. Konarev1, Galina S. Kachalova2, Alexandra Yu Ryazanova3, Elena A. Kubareva3, Anna S. Karyagina3,4,5, Hans D. Bartunik2, Dmitri I. Svergun1

1 European Molecular Biology Laboratory, Hamburg Outstation, Hamburg, Germany, 2Bach Institute of Biochemistry, Russian Academy of Sciences, Moscow, Russia, 3Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia, 4Gamaleya Institute of Epidemiology and Microbiology, Moscow, Russia, 5Institute of Agricultural Biotechnology, Russian Academy of Sciences, Moscow, Russia

Abstract

(Cytosine-5)-DNA methyltransferase SsoII (M.SsoII) consists of a methyltransferase domain (residues 72–379) and an N-terminal region (residues 1–71) which regulates transcription in SsoII restriction–modification system. Small-angle X-ray scattering (SAXS) is employed here to study the low resolution structure of M.SsoII and its complex with DNA containing the methylation site. The shapes reconstructed ab initio from the SAXS data reveal two distinct protein domains of unequal size. The larger domain matches the crystallographic structure of a homologous DNA methyltransferase Hhal (M.Hhal), and the cleft in this domain is occupied by DNA in the model of the complex reconstructed from the SAXS data. This larger domain can thus be identified as the methyltransferase domain whereas the other domain represents the N-terminal region. Homology modeling of the M.SsoII structure is performed by using the model of M.Hhal for the methyltransferase domain and representing the N-terminal region either as a flexible chain of dummy residues or as a rigid structure of a homologous protein (phage 434 repressor) connected to the methyltransferase domain by a short flexible linker. Both models are compatible with the SAXS data and demonstrate high mobility of the N-terminal region. The linker flexibility might play an important role in the function of M.SsoII as a transcription factor.

Citation: Konarev PV, Kachalova GS, Ryazanova AY, Kubareva EA, Karyagina AS, et al. (2014) Flexibility of the Linker between the Domains of DNA Methyltransferase SsoII Revealed by Small-Angle X-Ray Scattering: Implications for Transcription Regulation in SsoII Restriction–Modification System. PLoS ONE 9(4): e93453. doi:10.1371/journal.pone.0093453

Editor: Beata G. Vertessy, Institute of Enzymology of the Hungarian Academy of Science, Hungary

Received December 18, 2013; Accepted March 1, 2014; Published April 7, 2014

Copyright: © 2014 Konarev et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by Max-Planck-Gesellschaft zur Förderung der Wissenschaften and by Russian Foundation for Basic Research (grant nos. 13-04-00615_a and 12-04-32103). A.Yu.R. was a recipient of a Max-Planck fellowship. D.I.S. and P.V.K. acknowledge support from the German Ministry of Education and Science (BMBF) project BIOSCAT, Grant 05K20912. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: svergun@embl-hamburg.de

Introduction

DNA methyltransferases (MTases) catalyze methyl group transfer from S-adenosyl-L-methionine (AdoMet) to a certain base in DNA (cytosine or adenine). The cofactor AdoMet is converted into S-adenosyl-L-homocysteine (AdoHcy) in this reaction. Bacterial DNA MTases can be divided into the following 3 classes: (cytosine-5)-DNA MTases (C5-DNA MTases), N4-cytosine-DNA MTases, and N6-adenine-DNA MTases which methylate C5 atom of cytosine, N4 atom of cytosine, and N6 atom of adenine, respectively. Most known prokaryotic DNA MTases are components of Type II restriction-modification (R–M) systems which protect host cells from bacteriophage infection. A common Type II R–M system consists of a MTase which methylates certain DNA sequences and a restriction endonuclease (RE) which hydrolyses DNA if these sequences remain unmodified. An excessive RE activity can be dangerous for the host cell and the expression of the RE and MTase genes should thus be strictly coordinated [1]. To date, over 4000 R–M systems are characterized biochemically and/or genetically and over 20000 R–M systems are predicted bioinformatically (see REBASE) [2]. Among them, 3 variants of gene expression control at the transcriptional level are recognized: by a special C (controller) protein, by the MTase enzymatic activity, and by the MTase binding to a special regulatory site which differs from the methylation site [3].

The first method of gene expression control is based on the presence of a small gene encoding C-protein. This protein binds to an operator DNA sequence and regulates expression of its own gene as well as expression of the RE and MTase genes. Up to now, crystal structures of 4 different C-proteins are solved: C.AhdI [4], C.BclI [5], C.Csp231I [6], and C.Esp1396I [7]. All of them share highly similar three-dimensional structure including a classical helix-turn–helix (HTH) motif and are assigned to the Xre (xenobiotic response element) family of transcription regulators. All these C-proteins are dimeric in the crystallized form and C.AhdI has been shown to be a dimer in solution as well [8].

The second variant of gene expression control occurs for instance in the R–M system CfrBI. The MTase gene has a strong
promoter which overlaps with a weak promoter of the RE gene. A single methylation site, in turn, overlaps with the −35 promoter element of the MTase gene. Thus, the MTase enzymatic activity leads to methylation of the −35 element which provides the MTase gene repression and stimulates transcription of the RE gene [9].

The SsoII R–M system from Shigella sonnei has the third variant of gene expression control. The MTase of this R–M system, M.SsoII, is the main object of the present study. It belongs to C5-DNA MTases and methylates the second cytosine nucleotide (underlined) in the sequence 5′-CCNGG-3′/5′-GGNCC-5′ [10,11]. M.SsoII can also act as a transcription factor binding to a 15-bp quasipalindromic sequence 5′-AGGACAAAATGTCCCT-3′/5′-TCTGTTTACAGGA-5′ (the regulatory site) in the intergenic region of the SsoII R–M system and therefore downregulating the expression of its own gene and stimulating the expression of the cognate RE gene [12,13]. The same mechanism of action is shown for M.Ecl18kI [14], which differs from M.SsoII by a single amino acid residue. Some other C5-DNA MTases are shown experimentally to repress their own genes without any impact on expression of the corresponding REs, namely M.MspI [18], and M.ScrFI [19].

Sequence analysis demonstrates that M.SsoII contains 2 domains: a typical C5-DNA MTase domain (residues 72–379) and a regulatory domain (RD, residues 1–55). The latter one is predicted to contain an HTH motif [20,21] similarly to C-proteins and many other transcription regulators. To date, the Pfam database [22] contains 68 protein sequences which consist of a domain with the HTH motif followed by the C5-DNA MTase domain [23]. However, no crystallographic or NMR data about their structures are available as yet. The region between the RD and the MTase domain (residues 56–71) shares no similarity with any available high resolution model and contains 4 proline residues, suggesting this fragment to be potentially non-structured. The linker responsible for the interaction between the two domains of M.SsoII could play a crucial role in the functioning of M.SsoII in the cell. The N-terminal region (residues 1–71, i.e. the RD with the linker) determines the ability of M.SsoII to regulate transcription in the SsoII R–M system [12].

Since the mechanism of DNA methylation itself does not imply a dimer formation [24], most of DNA MTases exist in solution as monomers. On the contrary, transcription factors typically function as dimers and tetramers. Establishing the oligomeric state of M.SsoII in solution is therefore an important task with a clear functional implication. A deletion mutant representing only the MTase domain of M.SsoII has been found catalytically active but impossible to purify [25]. Because of this, another C5-DNA MTase, NlaX (M.NlaX), is used here as a control representing only the MTase domain. This enzyme shares 67% identity with the MTase domain of M.SsoII and has the same methylation specificity [11]. Due to the lack of additional domains, M.NlaX is transcriptionally inactive and can be regarded as a natural Δ(1–71) deletion mutant of M.SsoII.

In the present study, small-angle X-ray scattering (SAXS) is employed to determine the low resolution structures of apo-M.NlaX, apo-M.SsoII, and M.SsoII–15met complexed with a 15-bp DNA duplex containing the methylation site (15met). The obtained models of M.SsoII and M.NlaX are compared with the structure of M.HhaI, a one-domain C5-DNA MTase, which has been studied extensively by X-ray crystallography [26,27,28]. The SAXS data along with the results of size exclusion chromatography (SEC) and dynamic light scattering (DLS) unambiguously point to the monomeric state of apo-M.SsoII and of its complex with 15met even at higher solute concentrations. The low resolution model of full-length M.SsoII reveals an extended but folded structure of the N-terminal region as a distinct domain tethered by a highly flexible linker to the MTase domain. A possible role of the linker flexibility for transcription regulation in the SsoII R–M system is discussed.

### Materials and Methods

#### Protein expression and purification

E. coli strain M15 [pREP4] containing the plasmid pQMSsoII or pQMNlaX was grown at 37°C in LB medium with 30 μg/ml kanamycin and 50 μg/ml ampicillin to an A_{600} value of 0.6. Protein expression was induced with 0.7 mM isopropyl-β-D-galactopyranoside, and the cell culture was kept for 20 h at 20°C. The cells were harvested by centrifugation. The cell pellets were resuspended in buffer A (50 mM Na-phosphate, 100 mM NaCl, 5 mM β-mercaptoethanol, 5% (v/v) glycerol, pH 7.0) and
SAXS Reveals Linker Flexibility in M.SsoII

DNA–protein complex formation

DNA duplex 15met containing the M.SsoII methylation site was formed by annealing an equimolar mixture of 5’-AGAGCCAG-GAAGCAGGA-3’ and 5’-TCGGTTTCTGGCTTCT-3’ oligonucleotides (Metabion) in water, i.e. heating up to 70°C and cooling down slowly to room temperature. Complex formation between M.SsoII and 15met was carried out in buffer B (18 mM Tris-HCl, 136 mM NaCl, 3.5 mM β-mercaptoethanol, 10% (w/v) glycerol, pH 8.0) in the presence of AdoHcy, M.SsoII, AdoHcy, and the DNA duplex were mixed in ratio 1:2:1. The mixture was analysed by electrophoresis in 7% non-denaturing polyacrylamide gel. The gel was stained first with ethidium bromide (EtBr) to visualize DNA-containing bands and then with Coomassie Brilliant Blue to visualize protein-containing bands.

Size exclusion chromatography

Size exclusion chromatography was performed on a Superdex 75 10/300 column (GE Healthcare). The column was pre-equilibrated and eluted with buffer C (50 mM Tris-HCl, 100 mM NaCl, 5 mM β-mercaptoethanol, 5% (w/v) glycerol, pH 8.0). M.SsoII was loaded in concentration 3.2 mg/ml in buffer D (20 mM Tris-HCl, 100 mM NaCl, 5 mM β-mercaptoethanol, 20% (w/v) glycerol, pH 8.0). M.NlaX was loaded in concentration 7.7 mg/ml in buffer E (50 mM Tris-HCl, 100 mM NaCl, 5 mM β-mercaptoethanol, 20% (w/v) glycerol, pH 8.0). The column was previously calibrated with ribonuclease A (13700 Da), carbonic anhydrase (29000 Da), ovalbumin (monomer of 43000 Da and dimer of 86000 Da), and bovine serum albumin (BSA, monomer of 66000 Da) in buffer F (50 mM Tris-HCl, 200 mM NaCl, 5 mM β-mercaptoethanol, 5% (w/v) glycerol, pH 8.0). A NaCl concentration of 200 mM in buffer F served to estimate the total volume of the column by measuring conductivity. Blue dextran in buffer F was used to estimate the column void volume.

Dynamic light scattering

Dynamic light scattering measurements were performed using a ZetaSizer Nano-S (Malvern) with the laser wavelength of 633 nm in a quartz cuvette of 45 μl at 8°C. Samples of M.SsoII and M.NlaX were analyzed in 50 mM Na-phosphate buffer (pH 7.0) containing 5 mM β-mercaptoethanol with variations in glycerol or salt concentrations. The protein concentration was 0.5 mg/ml. The solution viscosities computed on the basis of glycerol concentration were 1.5217 cP and 1.8563 cP for water solutions containing 5% (w/v) and 15% (w/v) glycerol respectively. The column (GE Healthcare) and eluted with a 20–400 mM imidazole gradient. The target proteins were concentrated and their purity was estimated using 12.5% SDS-PAGE. Because of cytotoxicity of M.SsoII, the yield of purified M.SsoII was 0.14 mg from 1 l of cell culture, two orders of magnitude lower than that of M.NlaX.

Table 2. Overall parameters calculated from SAXS.

| Sample       | c, mg/ml | Rg, nm | Dmax, nm | Vp, nm³ | MMexp, kDa | ζab | ζn | ζom |
|--------------|----------|--------|----------|---------|------------|-----|----|-----|
| M.NlaX       | 1.6–7.0  | 2.36±0.04 | 7.0±0.5  | 68±7   | 33±4       | 1.12 | 1.21b | –   |
| M.SsoII      | 1.0–3.3  | 3.01±0.04 | 11.0±0.5 | 77±8   | 38±5       | 1.25 | 1.46 | 1.03 |
| M.SsoII–DNA  | 1.0–1.5  | 2.79±0.04 | 11.0±0.5 | 85±10  | 45±6       | 1.10 | 1.78 | 1.05 |

bNotations: Rg, radius of gyration; Dmax, maximum size of the particle; Vp, excluded volume of the hydrated particle; MMexp, experimental molecular mass of the solute; ζab, ζn and ζom, values for the fit from ab initio models, from rigid body modeling using BUNCH and from EOM, respectively.

Not in case of M.NlaX, ζab corresponds to the fit from the crystallographic structure of M.HhaI.

doi:10.1371/journal.pone.0093453.t002

doi:10.1371/journal.pone.0093453.g001

doi:10.1371/journal.pone.0093453.g002
SAXS Reveals Linker Flexibility in M.SsoII

range of concentrations appropriate to automatic choice of attenuation index was found to be 0.5–1.0 mg/ml for both proteins. The hydrodynamic diameter \( D_h \) was evaluated by the Stokes–Einstein equation from the autocorrelation function of the DLS measurements following standard procedures and the average MM was estimated using default Mark–Houwink parameters for a hard sphere.

SAXS measurements and data processing

Synchrotron radiation X-ray scattering data were collected on the EMBL X33 beamline at the DORIS III storage ring (DESY, Hamburg) [29]. Solutions of M.NlaX, M.SsoII, and M.SsoII–15met complex were measured for solute concentrations of 1.6–7.0 mg/ml, 1.0–3.3 mg/ml, and 1.0–1.5 mg/ml, respectively. A MAR345 image plate detector was used at the sample–detector distance 2.7 m and wavelength \( \lambda = 0.15 \) nm, covering the momentum transfer range 0.12\(<s<4.9 \) nm\(^{-1}\) \((s = 4\pi \sin\theta/\lambda, \) where \( 2\theta \) is the scattering angle). No radiation damage effects were detected by comparison of two data sets with 2-min exposure time. The data were averaged after normalization to the intensity of the incident beam, the scattering of the buffer was subtracted and the difference data were extrapolated to zero solute concentration using PRIMUS [30].

The radius of gyration \( R_g \) of solute protein molecule and the forward scattering \( I(0) \) were evaluated using the Guinier approximation at small angles \( (s < 1.3/R_g) \) assuming the intensity was represented as \( I(s) = I(0) \exp(-sR_g^2/3) \) and from the entire scattering pattern by the program GNOM [32]. In the latter case, the distance distribution functions \( p(r) \) and the maximum particle dimensions \( D_{max} \) were also computed. The molecular mass (MM) of the solute was evaluated by comparison of the calculated \( I(0) \) value with that of the standard solution of bovine serum albumin (MM of 66 kDa). The excluded volume of the hydrated protein molecule \( (V_p) \) was calculated using the Porod approximation:

\[
V_p = 2\pi^2 I(0) \int I_{exp}(s) ds
\]

in which the intensity \( I(s) \) was modified by subtraction of an appropriate constant from each data point to force the \( s^{-4} \) decay of the intensity at higher angles following Porod’s law [33] for homogeneous particles.

Low resolution \textit{ab initio} models of M.NlaX and M.SsoII were generated by DAMMIN [34], representing the protein by an assembly of densely packed beads. Simulated annealing was employed to build a compact interconnected configuration of beads inside a sphere with the diameter \( D_{max} \) that fits the experimental data \( I_{exp}(s) \) to minimize the discrepancy:

\[
\chi^2 = \frac{1}{N-1} \sum \frac{(I(s_j) - I_{calc}(s_j))^2}{\sigma(s_j)^2}
\]

where \( N \) is the number of experimental points, \( \epsilon \) is a scaling factor, \( I_{calc}(s_j) \) and \( \sigma(s_j) \) are the calculated intensity and the experimental error at the momentum transfer \( s_j \), respectively. The common structural features of a model were determined by averaging the configurations from ten separate runs using the program DAMAVER [35].

An alternative model of M.SsoII was constructed by homology modeling. The MTase domain of M.SsoII was represented by the crystallographic model of M.HhaI, a homologous one-domain MTase (PDB entry: 2HMY) [28], referred further as “M.HhaI model” (see Table 1). The scattering from M.HhaI was calculated using the program CRYSOL [36] and was compared with the experimental data for M.NlaX, which also consists of only one MTase domain. To construct the model of full-length M.SsoII, the N-terminal fragment was added to the M.HhaI structure by two alternative ways using BUNCH [37]. In the first case, referred further as “hybrid M.HhaI model”, the entire fragment (71 residues) was represented by an interconnected chain of dummy residues (DRs) [38]. In the second case (“hybrid M.HhaI-R434 model”), a crystallographic model of a protein homologous to the RD, phage 434 repressor (PDB entry: 1PER) [39] was used to represent the first 55 residues of the N-terminal fragment as a rigid
Figure 3. $R_g$ distributions from EOM for M.SsoII. Solid and dashed-dotted lines (curves 1 and 3) correspond to initial random pool; dashed and dotted lines (curves 2 and 4) to the selected ensembles (average of 50 independent EOM runs). Curves 1–2 correspond to the randomly generated N-terminal region of M.SsoII (“hybrid M.HhaI model”) and curves 3–4 were obtained with the “hybrid M.HhaI-R434 model”. The large width of the selected $R_g$ distributions supports the flexibility of the N-terminal region in the M.SsoII molecule (A). $R_g$ distributions from EOM for the M.SsoII–15met complex. The notations are the same as in (A) for the curves (1–2) (B).
doi:10.1371/journal.pone.0093453.g003

Results

Association state and overall parameters of M.SsoII and M.NlaX apo-forms in solution

The association states of M.NlaX and M.SsoII have first been studied by SEC (Figure S1) and DLS. The apparent MMs of the proteins have been estimated using the column calibration against the standard proteins set. The calculated values correspond to apparent MM of 34 kDa and 41 kDa for M.NlaX and M.SsoII, respectively.

In the DLS experiments, slight increase of PDI up to ~0.24 in case of M.NlaX could be prompted by a higher viscosity of the more concentrated M.NlaX solutions. The size distribution of M.SsoII contained a single narrow peak with polydispersity index (PDI) below 0.1. Estimated $D_h$ values varied in range 5.2–5.9 nm and 6.3–6.8 nm that corresponded to MMs of 32–43 kDa and 50–59 kDa for M.NlaX and M.SsoII, respectively. Note that the DLS calculations of MM are shape-dependent and this may lead to a somewhat overestimated MM of M.SsoII, which, as it will be seen later, is a rather elongated particle.

The information about MM has also been obtained from the SAXS experiments. The X-ray scattering intensity patterns $I(s)$ display no systematic changes with the solute concentration demonstrating no change in association state of M.NlaX and M.SsoII with concentration. The Guinier plots (initial portions of the scattering data in the coordinates $I \propto s^{-2}$) are linear.
suggesting homogeneity of the samples. The processed scattering data and the computed distance distribution functions are displayed in Figure 1. The overall parameters extracted from the SAXS data are summarized in Table 2.

The experimental MM of M.NlaX (33±4 kDa) suggests that the protein is monomeric in solution (theoretical MM of the monomer 50.3 kDa). This is further corroborated by the excluded volume $V_p$ of the particle 68±7 nm$^3$, in agreement with an empirical finding for globular proteins that the hydrated volume in nm$^3$ should numerically be about twice the MM in kDa. The experimental $R_g$ and $D_{max}$ (2.36±0.04 nm and 7.0±0.5 nm, respectively) point to a rather compact structure. The bell-shaped distance distribution function $p(r)$ for M.NlaX (Figure 1, insert) is also consistent with a compact shape of the protein.

The experimental MM of M.SsoII (30±5 kDa) and its $V_p$ (77±8 nm$^3$) indicate that this protein is also monomeric in solution (theoretical MM of the monomer 44.9 kDa). In contrast to M.NlaX, $R_g$ and $D_{max}$ values (3.01±0.04 nm and 11.0±0.5 nm, respectively) point to an elongated shape of M.SsoII and the $p(r)$ function for M.SsoII (Figure 1, insert) displays an asymmetric tail, typical for elongated particles.

Stoichiometry and overall parameters of the M.SsoII complex with the 15-bp DNA containing the methylation site

The DNA construct with a length of 15 bp has been chosen since M.SsoII methylation site must be flanked with at least 4 bps from each side for effective methylation [11]. M.SsoII has been mixed with 15met in the presence of AdoHcy, as AdoMet or AdoHcy presence is necessary for the specific complex formation between M.SsoII and its methylation site [42]. The resulting mixture has been analyzed by native gel electrophoresis. Coomassie staining indicates complex formation without an excess of unbound protein while EtBr staining demonstrates a minor band corresponding to a very small amount of unbound DNA (Figure S2). The SAXS analysis (Figure 1 and Table 2) yield the experimental MM (45±6 kDa) and $V_p$ (85±10 nm$^3$) pointing to a 1:1 stoichiometry for the M.SsoII–15met complex. Comparing the values of $R_g$ (2.79±0.04 nm) and $D_{max}$ (11.0±0.5 nm) for the complex with those for apo-M.SsoII, one can see that $D_{max}$ of the complex remains the same as that of M.SsoII, but the $R_g$ decreases, suggesting either the positioning of DNA in the central part of the complex or compaction of the entire structure. Still, the overall parameters of the M.SsoII–15met complex indicate an elongated shape and the $p(r)$ function of the complex displays an asymmetric tail (Figure 1, insert).

Two approaches for the molecule shape reconstruction

The macromolecular shapes have been reconstructed by two different approaches: ab initio modeling (using only the experimental X-ray scattering data) and hybrid rigid body modeling (using a crystallographic model of a homologous protein as a template). Ab initio low resolution models of M.NlaX (Figure 2A) and M.SsoII (Figure 2B) have been generated by DAMMIN [34] (see Methods for details). For the homology modeling, the atomic model of C5-DNA MTase HhaI (M.HhaI) has been chosen for which the crystal structure is available. M.HhaI methylates the inner cytosine residue in the sequence 5'-GGCC-3'/3'-CGGG-5', and the MTase domain of M.SsoII (as well as the whole M.NlaX) shares 41% identity with M.HhaI. The M.HhaI structure from PDB entry 2HMY [28] has been taken for modeling of M.NlaX and M.SsoII apo-forms. Both M.NlaX and M.HhaI consist of only the MTase domain. The full-length M.SsoII molecule contains in addition the N-terminal fragment, which was modeled (Figure 2B) in two alternative ways, as a DR-chain (“hybrid M.HhaI model”) or using a homologous structure [39], “hybrid M.HhaI-R434 model” (see model details in Methods).

The shape of the DNA–protein complex (Figure 2C) has also been independently reconstructed by ab initio and homology modeling. An ab initio model of the M.SsoII–15met complex has been generated by the program MONSA [40] (see Methods for details). For the homology modeling, the structure of M.HhaI in complex with the 15-bp DNA duplex containing its methylation site has been taken as a template (PDB entry: 3MHT) [27]. To reproduce full-length M.SsoII, the N-terminal fragment has been added to the M.HhaI structure by two alternative ways, similarly to the above construction of the apo-M.SsoII model (Table 1).

Ab initio shapes of M.NlaX, M.SsoII, and M.SsoII complex with the 15-bp DNA containing the methylation site

A typical low resolution shape of M.NlaX reconstructed ab initio (Figure 2A) has the overall size of about 7 nm x 4.4 nm x 3 nm and fits the experimental data with discrepancy $\chi = 1.1$ (Figure 1, curve 1, solid line). The scattering curve from the “M.HhaI model” calculated by CRYOSOL (see Methods) agrees with the experimental data ($\chi = 1.2$, Figure 1, curve 1, dashed line). This homology model overlaps well with the ab initio model (Figure 2A), suggesting that M.NlaX interacts with the shape close to that of the “M.HhaI model”.

The ab initio low resolution shape of M.SsoII is displayed in Figure 2B and fits the experimental data with $\chi = 1.3$ (Figure 1, curve 2, solid line). The model reveals two distinct domains, a “main” (larger) domain with the overall shape similar to that of M.NlaX, and an “additional” (smaller) domain presumably

Figure 4. Possible role of the linker flexibility in M.SsoII binding to target DNA. DNA is in grey, the methylation site is in cyan, the regulatory site is in red. The MTase domain is shown in blue. Apo-M.SsoII demonstrates high linker flexibility which results in coexistence of different conformations of the N-terminal region (shown by green and orange). Upon M.SsoII binding to the methylation site, the dominant RD conformers keep the same orientations as in the apo-form. Binding to the regulatory site is supposed to be a multi-stage process which results in M.SsoII dimer formation where both domains of each M.SsoII subunit are bound to the same DNA duplex. Such structural organization is likely to fix both domains in a certain position in relation to each other. As we do not know whether the N-terminal region position in the latter complex is similar to any of the dominant RD conformers in the M.SsoII complex with the methylation site, the N-terminal region is shown by another (yellow) color in the latter case.

doi:10.1371/journal.pone.0093453.g004
Flexibility of the N-terminal region in apo-M.SsoII and in the M.SsoII complex with the 15-bp DNA containing the methylation site

The presence of disordered portions in apo-M.SsoII and in the M.SsoII complex with the 15 bp DNA is qualitatively supported by the Kratky plots (Figure S4). These plots display broad bell-shaped peaks, with the scattering intensities multiplied by s² revealing upward trends at higher angles compared to the more downward trend observed for M.NlaX (where the N-term is missing). The increase of the higher angle portions of a Kratky plot in the M.SsoII complex with the 15-bp DNA is qualitatively supported also by the SAXS data (Figure S3) and reflect the flexibility of this region not only in the apo-M.SsoII but also in the M.SsoII–15met complex.

Rigid-body modeling of M.SsoII and M.SsoII complex with the 15-bp DNA containing the methylation site

To construct a more detailed model of M.SsoII, the “M.HhaI model” was treated as a rigid body, and the 71 N-terminal residues were represented as a chain of DRs yielding the “hybrid M.HhaI model”. Multiple runs of BUNCH (see Methods) starting from random initial configurations yield variable conformations of the N-terminal region, all providing good fits to the data with χ about 1.5. Some of the obtained models overlapped well with the ab initio model of M.SsoII (Figure 2B, black model), whereas the others displayed a tilted orientation of the N-terminal region with respect to the long axis of M.SsoII.

A similar approach was used to construct the model of the M.SsoII–15met complex. The M.HhaI complex with DNA has been treated as a rigid body and the missing N-terminal fragment of M.SsoII has been added by BUNCH, resulting in the “hybrid M.HhaI–DNA model”. Several BUNCH runs yielded an ensemble of solutions fitting the scattering data with χ about 1.8. Whereas the MTase domain of M.SsoII overlaps well with the larger domain of the ab initio model (Figure 2C), the presence of a variety of M.SsoII N-terminal region configurations suggests a significant flexibility of this region not only in the apo-M.SsoII but also in the M.SsoII–15met complex.

Discussion

In the earlier association state studies of DNA MTases in solution, diverging results have been reported. Some of DNA MTases are shown to exist predominantly as dimers, namely M.RsrI [43], M.MspI [44], and the Q237W mutant of M.HhaI [45], while some others remain monomeric, for example M.BamHI [46] and M.EcoRI [47]. In the present work, the oligomerization behavior of apo-M.SsoII was examined by SEC, DLS, and SAXS in a concentration range of 0.5–3.2 mg/ml. The data from all these methods are fully consistent with the monomeric state of apo-M.SsoII, and the protein remains monomeric upon binding to the 15met duplex.

Similar to most transcription factors bound to promoters comprising an inverted repeat [48,49,50], M.SsoII has to control its specificity and activity either by DNA-mediated oligomerization or by dimer self-assembly prior to the interaction with the promoter. Unfortunately, the M.SsoII complex with the regulatory site yielded a non-homogeneous reaction mixture and therefore could not be studied by SAXS. However, it has been shown recently that M.SsoII binds to a long DNA duplex (60-bp) with the regulatory site forming a complex with a stoichiometry protein:DNA = 2:1 [51]. No direct contacts between the protein subunits in the complex were observed, and, given the monomeric state of apo-M.SsoII in solution, it seems unlikely that the protein assemblies into dimers prior to the interaction with the regulatory site. It is therefore conceivable that DNA plays the major role in the formation of the M.SsoII complex with the regulatory site of the promoter region.

A typical C5-DNA MTase domain consists of 2 subdomains separated by a DNA-binding cleft. The larger subdomain comprises 10 motifs conservative for all C5-DNA MTases [52] and contains the AdoMet-binding site as well as the binding site for the target cytosine residue. The other subdomain (small, target recognition domain, TRD) carries a sequence, which is unique in every MTase and is responsible for the substrate specificity. Thus, the large subdomains of different C5-DNA MTases share high similarity in primary and tertiary structure while the small subdomains vary substantially in size and spatial structure [24].

The ab initio shapes of M.NlaX (Figure 2A) and apo-M.SsoII (Figure 2B) obtained from SAXS differ substantially from each other. The more elongated shape of M.SsoII demonstrates itself already in the noticeable increase of Rg and Dmax (Table 2) and in the asymmetric tail at the higher r of its p(r) distribution (Figure 1, insert). The M.SsoII head model displays two distinct domains, the
The “main” domain of M.SsoII in apo-form as well as in the complex with 15met matches the overall shape of M.NlaX and encompasses well the C$_2$-traces of M.HhaI conformations simulated for the template of the full-length M.SsoII. The prominent structural peculiarity of the “main” domain, its V-like cleft, is empty in the model of apo-M.SsoII and encloses the “DNA” beads in the M.SsoII–15met complex (Figure 2B–C). The mutual arrangement of “DNA” and “protein” beads is compatible with the MTase domain organization observed in the crystal structures of M.HhaI and M.HaeIII complexes with DNA [24,26,33], where the two protein subdomains embrace the DNA molecule. MTase binding to DNA containing its methylation site is known to mediate substantial conformational changes [26,33] leading to a more compact protein structure. The smaller $R_g$ value of the M.SsoII–15met complex model compared with the apo-M.SsoII strongly supports the identification of the “main” domain as a structural region corresponding to 72–379 residues of M.SsoII responsible for its methylation function.

The “additional” domain of the ab initio shape accommodates the N-terminal residues missing in the M.NlaX and M.HhaI sequences. The beads of this region are absent in the SAXS model of M.NlaX and the simulated conformations of M.HhaI also do not overlap with these beads. Sequence analysis of M.SsoII suggests that only a minor part of the N-terminal region is disordered while the major part (residues 1–55) represents a domain with a pronounced spatial structure. This assumption has been confirmed by circular dichroism spectroscopy combined with gel-shift assay [51]. A deletion mutant representing the N-terminal region of M.Eco10kI (differs from M.SsoII only by a single residue, Ile56Met) demonstrates a pronounced secondary structure and also retains the ability to bind specifically to the regulatory site, although with a lower affinity.

The models obtained with the “hybrid M.HhaI model” display a pronounced variability of the N-terminal fragment of M.SsoII (Figure 2B). The experimental scattering is also well described by rigid-body movements of the “hybrid M.HhaI-R134 model”, suggesting that the latter model is an adequate representation of the M.SsoII N-terminal region. Thus, the full-length M.SsoII can be described as a C$_5$-DNA MTase domain connected through a flexible linker to a folded RD which acts as a transcription regulator. This is represented schematically in Figure 4, where the “main” (MTase) domain of apo-M.SsoII is displayed in blue while the different possible orientations of the “additional” regulatory domain are depicted in green and orange.

The RD mobility is explored in more detail by the analysis of multiple coexistent conformers generated for the N-terminal residues using EOM. Comparison of the $R_g$ distributions calculated for the apo-M.SsoII (Figure 3A) and for the M.SsoII–15met complex (Figure 3B) reveals further details of the linker flexibility. Both distributions are bimodal and the $R_g$ values of both modes for the complex are smaller than those for the apo-form, in agreement with the observed decrease of the overall parameters of the M.SsoII–15met complex. Thus, the dominant RD conformers keep the same orientations in the complex as in the apo-form. Simultaneously, the two modes appear to be more distinct in the $R_g$ distribution for the complex, suggesting a somewhat more restricted conformational space for RD in the complex.

High linker flexibility in M.SsoII has recently been suggested on the basis of protein–protein crosslinking experiments [54], and the present work provides a direct structural evidence by a completely different technique, SAXS. The linker flexibility is likely to play an important role for the ability of M.SsoII to regulate transcription. This ability is based on M.SsoII binding to the regulatory site in the promoter region of the SsoII R–M system [12]. M.SsoII forms a stable complex with the regulatory site, which competes with RNA polymerase and therefore prevents transcription of the ssoIM gene [14,55]. This effect decreases the concentration of M.SsoII in the cell thus forming a regulatory circuit with a negative feedback. The ssoIR gene promoter is weaker than the ssoIM gene promoter and therefore repression of the ssoIM transcription stimulates the ssoIR gene transcription indirectly [14,55]. Thus, the regulatory activity of M.SsoII is in anticorrelation with its main function, DNA methylation. Switching between these two functions should be provided by M.SsoII binding either to the methylation site or to the regulatory site (Figure 4).

M.SsoII complex formation with the regulatory site is expected to be a multi-stage process (Figure 4). An unusual structure of the complex has been proposed on the basis of footprinting and crosslinking experiments: the M.SsoII N-terminal regions are bound to the regulatory site while the MTase domains are bound to DNA flanking the regulatory site (Figure 4) [54]. We suppose that the first step of the complex formation should be the RD binding to the regulatory site followed then by the MTase domain binding to the same DNA duplex in a non-specific manner which provides higher stability to the complex. In general, a high level of non-specific binding is typical for M.SsoII [56]. Such a structure where both M.SsoII domains are bound to the same DNA duplex is possible only in the case where the linker between the domains is extremely flexible. Indeed, its flexibility is confirmed in the present work. Since the catalytic centre in the M.SsoII complex with the regulatory site is occupied by non-specific DNA, M.SsoII can not bind to the methylation site anymore. Thus, the linker flexibility is a key structural feature which provides formation of the stable complex capable of transcription regulation and therefore switches off the methylation function of M.SsoII.
Figure S4 Kratky plots corresponding to the data in Figure 1. Experimental SAXS profiles were appropriately displaced along the logarithmic axis for better visualization. (TF)

Author Contributions
Conceived and designed the experiments: PVK GSK AYR EAK ASD HDB DIS. Performed the experiments: PVK GSK AYR. Analyzed the data: PVK GSK AYR DIS. Contributed reagents/materials/analysis tools: PVK GSK AYR EAK ASD HDB DIS. Wrote the paper: PVK GSK AYR EAK ASD HDB DIS.

References
1. Kobayashi I (2004) Restriction–modification systems as minimal forms of life. In: Pingoud A, editor. Restriction Endonucleases. Berlin: Springer-Verlag. pp. 19–60.
2. Roberts RJ, Vincze T, Posfai J, Macelis D (2010) REBASE: a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res 38: D234–236.
3. Nagornykh MO, Bogdanova ES, Protosenko AS, Zakharova MV, Solonin AS, et al. (2008) Regulation of gene expression in type II restriction-modification system. Genetika (Moscow) 44: 606–613.
4. McGeehan JE, Streeter SD, Papapanagiotou I, Fox GC, Kneale GG (2003) High-resolution crystal structure of the restriction-modification controller protein C.AhdI from Aeromonas hydrophila. J Mol Biol 346: 689–701.
5. Sawaya MR, Zhu Z, Mersha F, Chan SH, Dabur R, et al. (2005) Crystal structure of the restriction-modification system control element C.Bcl and its cognate binding site. Structure 13: 1365–1372.
6. McGeehan JE, Streeter SD, Threlf SJ, Taylor JE, Shevstov MB, et al. (2011) Structural analysis of a novel class of R-M controller proteins: C.Caps11 from Citrobacter sp. RFL231. J Mol Biol 409: 179–188.
7. Ball N, Streeter SD, Kneale GG, McGeehan JE (2009) Structure of the restriction-modification controller protein C.Equil1. Acta Crystallogr D Biol Crystallogr 65: 900–903.
8. Streeter SD, Papapanagiotou I, McGeehan JE, Kneale GG (2004) DNA footprinting and biophysical characterization of the controller protein C.AhdI suggests the basis of a genetic switch. Nucleic Acids Res 32: 6445–6453.
9. Beletskaya IV, Zakharova MV, Shlyapnikov MG, Semenova LM, Solonin AS (2000) DNA methylation at the C.Bcl site is involved in expression control in the C.Bcl restriction-modification system. Nucleic Acids Res 28: 3817–3822.
10. Nikolaykina LI, Topagni NA, Suchkov SV, Kartashova IM, Debok SS (1984) Sequence specificity of isolated DNA-cytosine methylases from Shigella sonnei 47 cells. Biochem Int 9: 771–781.
11. Kubareva EA, Walter J, Karyagina AS, Vorob’eva OV, Lui PC, et al. (2002) Determination of methylation site of DNA-methyltransferase NaX by a hybrid method. Biotechniques 33: 526–331.
12. Karyagina AS, Shih I, Tashlitskii V, Khodoun M, Vas’iev S, et al. (1997) Specific binding of sso II DNA methyltransferase to its promoter region provides the regulation of sso II restriction-modification gene expression. Nucleic Acids Res 25: 2114–2120.
13. Shih I, Tashlitskii V, Khodoun M, Vas’iev S, Alekseev Y, et al. (1998) DNA-methylation SsoII interaction with own promoter region binding site. Nucleic Acids Res 26: 2659–2664.
14. Protosenko A, Zakharova M, Nagornykh M, Solonin A, Severinov K (2009) Transcription repression of restriction-modification system C.Equil1. Nucleic Acids Res 37: 5322–5330.
15. Som S, Friedman S (1995) Autonomous regulation of the EcoRII methylase gene at the transcriptional level: effect of 5-aza-cytosine. EMBO J 14: 4297–4303.
16. Friedman S, Som S (1993) Induction of EcoRII methylase: evidence for autonomous control. J Bacteriol 175: 6293–6298.
17. O’Gara M, Zhang X, Roberts RJ (1999) Structure of the LlaJI restriction/modification system. Mol Biol (Mosk) 34: 921–926.
18. O’Gara M, Zhang X, Roberts RJ, Cheng X (1994) HhaI methyltransferase flips its target base out of the DNA helix. Cell 76: 357–369.
19. O’Gara M, Zhang X, Roberts RJ, Cheng X (1996) Enzymatic C5 cytosine-methylation of DNA: mechanistic implications of new crystal structures for HhaI methyltransferase-DNA-AdoHcy complexes. J Mol Biol 261: 641–645.
20. O’Gara M, Zhang X, Roberts RJ, Cheng X (1999) Structure of a binary complex of HhaI methyltransferase with Sadenosyl-L-methionine formed in the presence of a short non-specific DNA oligonucleotide. J Mol Biol 287: 201–209.
21. Roelse MV, Klaring R, Ristau U, Kohrblain B, Jahn D, et al. (2007) Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg, Germany. J Appl Crystallogr 40: 1579–1583.
22. von Einsiedel B, Neubert A, Bickel J, Schulze A, et al. (2010) Secondary structure of SsoII-like (cytosine-5)-DNA methyltransferase SoolI with target and regulatory DNAs. Biopolymers 57: 1532–1544.
23. Vorob’eva OV, Karyagina AS, Volkov EM, Rolaev MB, Oretskaya TS, et al. (2010) Secondary structure of sso II DNA methyltransferase. Nucleic Acids Res 38: D234–D242.
24. Bohmke R, Lefebvre M, Rupley D, Grimaldi C, Dromey C, et al. (1993) Crystallographic structure of the HhaI methyltransferase. J Mol Biol 233: 1045–1057.
25. Dong A, Zhou L, Zhang X, Stichel S, Roberts RJ, et al. (2004) Structure of the Q537W mutant of HhaI DNA methyltransferase: an insight into protein-protein interactions. Biocell 28: 373–379.
26. Nardone G, George J, Chirikjian GJ (1984) Sequence-specific BamHI methylase. Purification and characterization. J Biol Chem 259: 10357–10362.
27. Rubin RA, Modrich P (1977) EcoRI methylase. Physical and catalytic properties of autogenous control. J Bacteriol 137: 6293–6298.
28. O’Gara M, Zhang X, Roberts RJ, Cheng X (1996) HhaI methyltransferase flips its target base out of the DNA helix. Cell 76: 357–369.
29. Rodger DW, Harrison SC (1993) The complex between phage 434 repressor and DNA. J Appl Crystallogr 26: 120–127.
30. Svergun DI, Petoukhov MV, Koehl MJH (2001) Determination of domain structure of proteins from X-ray scattering data. Biophys J 80: 2946–2953.
31. Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 76: 2872–2886.
32. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small angle scattering. J Appl Crystallogr 36: 860–864.
33. Svergun DI, Barberato C, Koehl MJH (1995) CRYOSOL - a program to evaluate X-ray solution scattering of biocomplex macromolecules from atomic coordinates. J Appl Crystallogr 28: 767–771.
34. Petoukhov MV, Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys J 89: 1237–1250.
35. Svergun DI, Petoukhov MV, Koehl MJH (2001) Determination of domain structure of proteins from X-ray scattering data. Biophys J 80: 2946–2953.
36. Rodgers DW, Harrison SC (1993) The complex between phage 434 repressor DNA-binding domain and operator site OR3: structural differences between consensus and non-consensus half-sites. Structure 1: 227–240.
37. Svergun DI, Nierhaus KH (2003) Analysis of protein-RNA distribution in the 70 S Escherichia coli ribosome. J Biol Chem 278: 14432–14439.
38. Bernaldo P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc 129: 5656–5664.
39. Vorob’eva OV, Karyagina AS, Volkov EM, Virovets MB, Oretskaya TS, et al. (2002) An analysis of methyltransferase SoolI DNA contacts in the enzyme–substrate complex. Russian J Bioorg Chem 28: 303–370.
40. Thomas CB, Gumpert RJ (2006) Dimethylation of the bacterial Rif N6-adenine DNA methyltransferase. Nucleic Acids Res 34: 297–307.
41. Dubey AK, Mollet B, Roberts RJ (1992) Purification and characterization of the SoolI DNA methyltransferase cloned and overexpressed in E. coli. Nucleic Acids Res 20: 1579–1583.
42. Dong A, Zhou L, Zhang X, Stichel S, Roberts RJ, et al. (2004) Structure of the Q537W mutant of HhaI DNA methyltransferase: an insight into protein-protein interactions. Biocell 28: 373–379.
54. Ryazanova AY, Winkler I, Friedhoff P, Viryasov MB, Oretskaya TS, et al. (2011) Crosslinking of (cytosine-5)-DNA methyltransferase SsoII and its complexes with specific DNA duplexes provides an insight into their structures. Nucleosides Nucleotides Nucleic Acids 30: 632–650.

55. Burenina OY, Fedotova EA, Ryazanova AY, Protenko AS, Zakharova MV, et al. (2013) Peculiarities of the regulation of gene expression in the Ec118kI restriction-modification system. Acta Naturae 5: 70–80.

56. Ryazanova AY, Kubareva EA, Grman I, Lavrova NV, Ryazanova EM, et al. (2011) The study of the interaction of (cytosine-5)-DNA methyltransferase SsoII with DNA by acoustic method. Analyst 136: 1227–1233.