The structure of M.EcoKI Type I DNA methyltransferase with a DNA mimic antirestriction protein

Christopher K. Kennaway1, Agnieszka Obarska-Kosinska2,3, John H. White4, Irina Tusznyska2,3, Laurie P. Cooper4, Janusz M. Bujnicki2,5, John Trinick1 and David T. F. Dryden4,*

1Astbury Centre, Institute of Molecular and Cellular Biology, University of Leeds, UK, 2Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, Trojdena 4, PL-02-109 Warsaw, 3Institute of Biochemistry and Biophysics PAS, Pawinskiego 5A, 02-106 Warsaw, Poland, 4School of Chemistry, University of Edinburgh, The Kings’ Buildings, Edinburgh, EH9 3JJ, UK and 5Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Umultowska 89, PL-61-614 Poznan, Poland

Received October 16, 2008; Revised November 20, 2008; Accepted November 21, 2008

ABSTRACT

Type-I DNA restriction–modification (R/M) systems are important agents in limiting the transmission of mobile genetic elements responsible for spreading bacterial resistance to antibiotics. EcoKI, a Type I R/M enzyme from Escherichia coli, acts by methylation- and sequence-specific recognition, leading to either methylation of DNA or translocation and cutting at a random site, often hundreds of base pairs away. Consisting of one specificity subunit, two modification subunits, and two DNA translocase/endonuclease subunits, EcoKI is inhibited by the T7 phage antirestriction protein ocr, a DNA mimic. We present a 3D density map generated by negative-stain electron microscopy and single particle analysis of the central core of the restriction complex, the M.EcoKI M2S1 methyltransferase, bound to ocr. We also present complete atomic models of M.EcoKI in complex with ocr and its cognate DNA giving a clear picture of the overall clamp-like operation of the enzyme. The model is consistent with a large body of experimental data on EcoKI published over 40 years.

INTRODUCTION

The EcoKI Type I DNA restriction–modification (R/M) enzyme was the first R/M enzyme to be discovered (1) and purified (2). It was determined that EcoKI, like all Type I R/M enzymes (3–8), is a complex oligomer formed from a core methyltransferase (MTase, M.EcoKI) comprised one HsdS DNA Specificity subunit and two HsdM DNA Modification subunits, M2S1, with a total molecular weight of ∼169 kDa. This is complexed with two additional HsdR DNA Restriction subunits, giving a complex R2M2S1 with a total molecular weight of ∼440 kDa that is capable of functioning as both MTase and restriction endonuclease (EcoKI or R.EcoKI) (9,10). Other Type I R/M enzymes display more variable subunit structures with some subunits weakly bound or capable of forming larger aggregates (11–13). Nevertheless, the core functions are performed by the M2S1 MTase and the R2M2S1 bifunctional MTase/endonuclease. Despite much subsequent work, it is only recently that crystal structures of HsdS from Methanococcus jannaschii and Mycoplasma genitalium (14,15) and HsdM from Escherichia coli and Bacteroides thetaiotaomicron (pdb code: 2ar0, New York Structural GenomiX Research Consortium, DOI 10.2210/pdb2ar0/pdb and pdb code: 2okc, Joint Center for Structural Genomics, DOI 10.2210/pdb2okc/pdb, respectively) have become available. The overall structure is not known for either the MTase or the complete R/M complex although several models have been proposed (16–20).

The absence of these structures has hindered study of Type I R/M enzymes but their enduring appeal as research targets (21) is due to their extraordinary complexity of operation. This comprises recognition of a bipartite, asymmetric DNA specificity sequence (e.g. EcoKI recog-
nizes $\text{AAC}[\text{N}_6\text{GTGC}]$, the recognition of methylation of the adenines at the bold underlined positions, and switching between an MTase activity on newly replicated, hemi-methylated host DNA and an endonuclease activity on unmethylated foreign DNA (3). The latter reaction requires massive ATP hydrolysis to drive translocation of up to 50 kb of DNA (22,23) at rates of up to 1 kb per second (24–29). The DNA is then cleaved at a random sequence remote from the original specificity sequence, with the generation of variable length single-strand overhangs (28,29). DNA-induced dimerization of EcoKI prior to translocation has been observed using atomic force microscopy (30,31). These varied activities have functional analogies with the operation of other complex DNA-manipulating machines (32,33). Of further note is the presence of two target recognition domains (TRDs) in HsdS that can be easily exchanged for other TRDs to generate novel DNA sequence specificities. The Type I R/M systems can be divided into, at present, five related families, IA to IE, defined by genetic complementation, antibody cross reactivity, DNA hybridization and sequence comparisons (34,35).

Of great topical interest due to the spread of drug resistance by horizontal gene transfer in bacteria (36,37), is the susceptibility of the R/M enzymes to antirestriction measures encoded by a variety of mobile genetic elements such as phage, conjugative plasmids and transposons (38–43). These antirestriction measures include the production of proteins that structurally mimic DNA (44). DNA mimics block the binding sites within R/M enzymes and inactivate them (45–47). The dimeric protein ocr from $\text{Tha} \phi$ phage matches the shape and surface potential characteristics of double-stranded DNA bent to the degree necessary for recognition by EcoKI.

We now present a model of M.EcoKI bound to ocr at $\sim 18 \text{ Å}$ resolution as determined by negative-stain electron microscopy (EM). Combining the EM model with new atomic models generated from the known crystallographic structures of the subunits allows a detailed structure of the M.EcoKI enzyme to be constructed. This structure is consistent with and rationalizes many experimental results obtained with M.EcoKI and other Type I MTases.

**MATERIALS AND METHODS**

**Protein samples**

M.EcoKI and the ocr protein were prepared as previously described (9,10,46). A complex of M.EcoKI and ocr was prepared by mixing the two proteins together at a molar ratio of 1:1.5 (29 $\mu$M M.EcoKI and 44 $\mu$M dimeric ocr) in 20 mM Tris, 100 mM NaCl, pH 7 buffer and incubating at room temperature for 5 min. The excess of ocr was not visible in the EM due to its small size.

**EM**

Grids were made by placing M.EcoKI and ocr complexes diluted 400 times onto UV-treated (48) continuous carbon-coated copper mesh grids, then stained with 1% uranyl acetate solution. Grids were viewed in a Jeol 1200EX electron microscope fitted with a LaB$_6$ electron source operating at 80 kV in low dose mode. Negatives (Kodak SO63) were taken at either 40 000× or 50 000× magnification with defocus ranging from 250 to 850 nm, and digitized at 6.25 μm step size using a Coolscan8000 (Nikon). The defocus and astigmatism were determined with CTFIND3 (49), then 17807 particles were picked from selected micrographs using Boxer (EMAN) (50) in autobox mode. Particles were band-pass filtered using Imagic (ImageScience, Germany), with low frequencies below 1/150 Å and those above the first zero of the contrast transfer function (CTF) being removed. After interpolating the 40 000× and 50 000× magnified particles to a common scale (3.125 Å/pixel), initial class average images without reference bias were generated in Imagic (51) by multi-variate statistical analysis and classification.

Six independent starting models consisting of 16 randomly placed Gaussian spheres were made in EMAN and each used in nine rounds of 3D refinement with the dataset. Two of the randomly generated starting models converged on a map similar to others made with some degree of *a priori* knowledge of the structure using the angular reconstitution method (52). These two 3D maps also showed approximate 2-fold symmetry. One was taken for further refinements in EMAN with C2 point group symmetry imposed. At several steps during refinement, classes that had the best match to the reprojections were selected and new 3D maps made. The final map has a nominal resolution of 18 Å at 0.5 Fourier shell cross-correlation using the odd–even test, and was 3D Fourier filtered at 12 Å. Maps were visualized using UCSF Chimera (53).

**Model building**

Full details of the model building are described in the Supplementary Data and so are described only briefly here. The model of M.EcoKI bound to DNA or to ocr was built using the following crystal structures: HsdS(MjaXIP) from *M. jannaschii* (pdb code: 1yl2) (15), HsdM (EcoKI) (pdb code: 2ar0, New York Structural GenomiX Research Consortium, DOI 10.2210/pdb2ar0/pdb) and HsdM from *B. thetaiaotaomicron* VPI-5482 (pdb code: 20kc, Joint Center for Structural Genomics, DOI 10.2210/pdb20kc/pdb), the structure of a Type II MTase bound to its cognate DNA duplex, M.TaqI–DNA (pdb code: 1g38) (54), and the ocr structure (pdb code: 1s7z) (44). The M.TaqI–DNA structure was used for modelling of the HsdS–HsdM orientation. HsdS was modelled on the 1yf2 template by homology modelling of the crystal structure of a related HsdM from *B. thetaiaotaomicron* VPI-5482 without reference bias were generated in Imagic (51) by multi-variate statistical analysis and classification. Six independent starting models converged on a map similar to others made with some degree of *a priori* knowledge of the structure using the angular reconstitution method (52). These two 3D maps also showed approximate 2-fold symmetry. One was taken for further refinements in EMAN with C2 point group symmetry imposed. At several steps during refinement, classes that had the best match to the reprojections were selected and new 3D maps made. The final map has a nominal resolution of 18 Å at 0.5 Fourier shell cross-correlation using the odd–even test, and was 3D Fourier filtered at 12 Å. Maps were visualized using UCSF Chimera (53).

To build the M.EcoKI–DNA model, DNA was added to the HsdS model based on the orientation of DNA in 1g38. HsdM were docked together using ROSETTA (55). The crystal structure of the EcoKI HsdM has been solved, but we have remodelled its domain–domain contacts based on the crystal structure of a related HsdM from *B. thetaiaotaomicron* VPI-5482. The missing C-terminal regions of EcoKI HsdM were modelled *de novo* using ROSETTA and fitted into the EM map for M.EcoKI–ocr described below.
HsdS–DNA model was combined with the HsdM dimer model so as to satisfy the orientation between the TRD and the catalytic domains of M.TaqI. To find the best fits, normal mode analysis was performed to generate a range of conformations of HsdS and HsdM.

The M.EcoK1–ocr model was constructed by multi-body flexible docking using HADDOCK (56). During the docking procedure, residues of HsdS and HsdM that form contacts with the DNA in the model of M.EcoK1–DNA were restrained to make contact with ocr, based on the assumption that ocr binds to HsdS and HsdM in a similar manner as DNA (44). The alternative models of M.EcoK1–ocr were fitted to the EM map as single rigid bodies using the COLORES module from the SITUS package (57) and the best scoring model was selected.

RESULTS

EM results

Addition of stoichiometric quantities of the ocr inhibitor was found to stabilize the otherwise labile M.EcoK1 complex. Individual molecules of M.EcoK1–ocr complexes embedded in heavy metal stain were imaged by EM and analysed by single particle methods. Particles appeared to be homogeneously sized and were well distributed. Class averages showed a wide variety of views and exhibited considerable detail (Figure 1a) that allowed a unique 3D reconstruction of the complex to a nominal resolution of 18 Å (Figure 1b and c). The atomic model described below has been placed within the EM reconstruction in Figure 1b and has been used to colour the reconstruction surface in Figure 1c. Although at the atomic level, the complex is asymmetric, 2-fold symmetry was imposed since HsdS is known to have strong pseudo-symmetry (14,15,17). The EM map longest dimension was ~120 Å and 100 Å along the base, and the overall shape was an excellent match to the complex modelled in silico based on biochemical and X-ray data, as described below. As described below and in the Supplementary Data, when one takes the size, shape and symmetry constraints of the HsdS, HsdM and ocr crystal structures and constraints from biochemical data, only one arrangement of these structures within the EM reconstruction was possible. Given this unique solution, density attributable to the two-helix coiled-coil spacer of HsdS could be clearly identified, as well as regions attributable to the two HsdM. In addition to contacting the TRDs of HsdS, the two HsdM wrap around ocr and make contact with each other. An extension projects from the bottom of each HsdM to either side of the

Figure 1. EM data, 3D reconstructions and model for the M.EcoK1–ocr complex. (a) Eight selected EM class average images (top row) with their corresponding reprojections of the EM map (second row), showing a range of views. (b) Two orthogonal semi-transparent surface representations of the EM 3D reconstruction each with a view of the modelled coordinates fitted as a rigid body (green and red—HsdM; blue—ocr dimer; yellow—HsdS). (c) Three surface views of the M.EcoK1–ocr EM 3D reconstruction coloured according to proximity to the fitted coordinates, with the positions of individual protein chains indicated with arrows. (d) A stereo view of the M.EcoK1–ocr atomic model. (e) Schematic diagram of proposed mechanism of clamping and release of DNA substrate (light blue) facilitated by twisting of the coiled-coil (yellow, viewed end-on).
HsdS coiled-coil spacer. This density returned in subsequent 3D refinements if masked out of the EM map, supporting the presence of these regions. These extensions could be attributed to the C-terminal regions of HsdM that are not present in the X-ray coordinates, as described below. Density in the centre of the map suggested the presence of the bound inhibitor protein ocr as it is known that ocr is fully enfolded by HsdS and HsdM (47). Sufficient detail was not present to show its exact orientation but its general position could be inferred. Features in the 3D map such as the straight coiled-coil spacer, which was ~16 Å wide, could also be seen in many class average images (Figure 1a). That features of this size were visible, just below the limit of resolution as measured by FSC, attests to the quality of the map and suggests that the 18 Å estimate of resolution is conservative.

**Model of the M.EcoKI–ocr complex**

The general shape of the model of M.EcoKI in complex with ocr dimer is shown in Figure 1d. This atomic model was constructed independently of the EM imaging based upon crystal structures of subunits and biochemical data. The MTase forms a clamp-like structure encircling the ocr dimer. The coiled-coil region of HsdS contains a considerable kink to allow the ocr molecule to fit optimally within this clamp. Each TRD of HsdS contacts one monomer of the ocr molecule and one HsdM. The exact location of the C-terminal helix of the HsdS subunit is speculative but has little effect on the overall outline shape of the molecule. The C-terminal regions of each HsdM, aa 470–529, project down to contact the coiled-coil region of HsdS. The location of the C-terminal regions of each HsdM is also speculative but the postulated location does bring conserved regions of HsdM and HsdS into close contact as discussed later and does fit with the EM model. This is important as the other region of contact between HsdM and HsdS uses the TRD regions of HsdS; these are highly variable in sequence in Type I R/M enzymes and not expected to provide a good HsdM–HsdS interface. The interface between the two HsdM is formed by their N-terminal domains, aa 1–153. The clamp-like shape of the MTase model suggests that it must open up along the interface between the two HsdM to allow access for ocr or DNA (Figure 1e). The interfaces between the subunits in the MTase–ocr complex are less extensive than in the MTase–DNA complex described below and the whole MTase appears to be strained to accommodate the ocr molecule that is slightly larger than a DNA duplex. The model structure with ocr bound may represent a ‘partially open clamp’ state, while the model with DNA bound would be a ‘fully closed clamp’ state. Opening of the clamp to allow ocr (or DNA) binding may be accomplished by simply flexing the coiled-coil region of HsdS. Normal mode calculations indicate that such motions should occur with sufficient amplitude to allow binding of ocr or DNA (data not shown). Having constructed the atomic M.EcoKI–ocr model, it is obvious that this can be simply docked into the EM model as shown (Figure 1b). A comparison of the fit of the EcoKI MTase–ocr complex model to the EM map with the fits obtained for models in which subunit structures are randomly docked in to the EM map indicate that our data driven model fits the EM map significantly better than random models (see Supplementary Data).

**Model of the M.EcoKI–DNA complex**

The model of M.EcoKI bound to DNA is shown in Figure 2. The two major differences between this and the partial models of Type I R/M enzymes published previously (16–19) are that the locations of the N-terminal and C-terminal regions of the HsdM are now defined within the overall structure. Most importantly, the N-terminal domains of HsdM can be seen to form a bridge over the DNA duplex, thus forming a clamp-like structure encircling the DNA, which is a key feature that was missing in earlier models. The location of the C-terminal regions of each HsdM with respect to HsdS is slightly different from that in M.EcoKI–ocr model, since the relative orientation of the subunits is altered. The coiled-coil region linking the two TRDs is straighter and the subunit interfaces within the MTase are more extensive in area than in the M.EcoKI–ocr model as the DNA duplex appears to fit better within the MTase allowing the clamp to close fully. The DNA bound to the TRDs contains a sharp bend in the non-specific spacer part of the target sequence and the bend angle of ~45° matches that observed experimentally (44,58).

**DISCUSSION**

The proposed *in silico* models, based upon crystallographic data, fit extremely well with the EM reconstruction and are attractive atomic resolution solutions to the structure of the MTase core of a Type I R/M enzyme. It is important to assess whether the models can accommodate the large body of experimental information concerning M.EcoKI. As most of the published work concerns interaction between EcoKI and DNA, we focus the following...
discussion primarily on the M.EcoKI–DNA model rather than the M.EcoKI–ocr model. Each section of the discussion describes how experimental data can be accommodated by the EM structure and the derived structural model and indicates new regions of interaction, which could be investigated in the future. Lastly, we compare the structural model of M.EcoKI with those recently proposed for the type IC enzyme, EcoR124I (19,20).

Analysis of the M.EcoKI–DNA model

Interface between the two HsdM via their N-terminal domains. Our model fits each HsdM into the EM structure in such a way as to bring their N-terminal domains into contact with each other (Figure 3a). They effectively form a bridge over the DNA and this bridge must swing open to allow the DNA to enter and be recognized by
HsdS. It is known that the ocr protein, which acts as a DNA mimic, is encircled by M.EcoKI (47) supporting the positioning of the N-terminal domains to complete the clamp encircling the DNA. The main interaction between these two domains is via helix G_{64}QELQFYRK MVHL_{78} contacting, in an antiparallel manner, the equivalent helix within the other HsdM. A second contact region is between a loop H_{91}NVSTTTI_{98} with the equivalent loop in the other subunit. The helical axes and loops in this interface run approximately parallel to the axis of the coiled-coil in HsdS.

The N-terminal domain of each HsdM contains the m* mutations (Figure 3a–c) (59), so called because they change the strong preference of M.EcoKI for methylated hemimethylated DNA target sequences (9) to an ability to additionally methylate unmethylated targets. It is noteworthy that the disruption caused by the m* mutations does not affect the ability of the enzyme to methylate DNA, only its ability to recognize the methylation state of the target sequence. A single m* mutation, R72Q, lies within the predicted helix G_{64}QELQFYRK MVHL_{78} and contacts the equivalent helix in the other HsdM (Figure 3a). This mutation has an r+m* phenotype suggesting a small deformation of the interface affecting communication of the methylation status of the two adenines in the target sequence and closure of the proposed clamp around the DNA but without affecting restriction by a complete EcoKI.

The mutations Q48K, L85Q, R124L, R124C, L134V, R153H and R153S also give rise to an r+m* phenotype. These locations in the model lie on the interface between the N-terminal m* domain (aa 1–153) and the middle catalytic domain (aa 154–469) of the same HsdM. It is possible to envisage how these non-conservative changes will disrupt this interface resulting in the observed r+m* phenotype, perhaps by interfering with closure of the clamp.

The remaining mutations, L113R, L113P, L113Q, W115R and S144Y all give an r+m* phenotype. These appear to be silent. Early modelling studies (64) suggested that these were not on the protein–DNA interface. The observation that restriction activity is lost because of these mutations may further suggest that this small region interacts with the HsdR subunit in the complete EcoKI.

Interactions between HsdM and HsdS. The EM density at the bottom of HsdM has a protrusion that passes close to the HsdS coiled-coil and we believe this to be the C-terminal region of HsdM comprising residues 470–529. This region is disordered in the two HsdM crystal structures available suggesting that it is flexible in the absence of HsdS. The EM results support the rigidity of the extensions in the complex, since flexible areas are invariably blurred out by the averaging inherent in the single particle methodology. However, the limited resolution of the EM map, and lack of X-ray data, only allows the computational atomic model to be positioned approximately.

Sequence analysis indicates that the C-terminal region of HsdM has a propensity to form α-helices and the model suggests that these reach out and contact the coiled-coil in HsdS, as shown in Figure 3d. These contacts would play a central role in the assembly of the MTase providing an interface between known regions of high sequence conservation in both HsdS and HsdM of the Type IA family (5,8). This high degree of conservation in the interface between HsdS and HsdM would be essential to facilitate the exchange of subunits of different Type IA family systems by complementation. Experimental support for this aspect of the EM and computational model comes from proteolysis experiments. It was found that ~43 residues from the C-terminal region of HsdM could be easily removed, leaving a stable, 55-kDa folded protein (60,61). However, this truncated HsdM was unable to bind to M1_S complexes indicating that a subunit interface had been removed (60).

In the model, a second HsdS–HsdM interface occurs via a helix-loop structure, N_{133}KISSLSAGANIA_{144}, within the first TRD of HsdS and two loops, aa 430–433 and aa 464–470, on the edge of the catalytic domain within HsdM (not shown in Figure 3). Contacting the same region of the second HsdM, one can identify an equivalent helix and loop in the second TRD of HsdS as A_{346}MMNCVKTTSQK_{358}. It is known that the TRDs within a family of Type I R/M systems can be exchanged easily to generate new sequence specificities (5,8). However, the TRD amino acid sequences are very variable and would not be expected to offer enough conservation to allow the formation of a particularly stable interface with HsdM. Hence, these regions must be of secondary importance to the regions discussed in the previous paragraph.

HsdS to DNA contacts. The N-terminal part of HsdS forms the first TRD, which recognizes the sequence 5'-AAC-3'. Many mutations in this TRD of HsdS from EcoKI have been catalogued (62,63) but the majority of these were silent. Early modelling studies (64) suggested that these were not on the protein–DNA interface. However, two loops were identified in this analysis, aa 83–91 and aa 96–118, as contacting the DNA and this is corroborated in the current model. Within the second of these loops, mutations of aa 91, 95, 105, 107 and 108 led to a loss of activity although the protein was folded and could bind DNA (62,63). In our model, these appear to contact the complementary strand to the 5'-AAC-3' sequence (Figure 3e). Chen et al. (65) identified a further amino acid, Y27, by cross-linking experiments in contact with the DNA target. In our map, Y27 is within 4 Å of the minor groove of DNA (not visible in Figure 3e as it is behind the DNA).

In the first TRD of HsdS, our model has an additional loop to those identified previously, S_{139}AGANINNIK_{148} (Figure 3e). This loop is closest to the major groove DNA bases forming the target sequence 5'-AAC-3'. Mutations S139P, G141V and G141A affected activity but other more conservative substitutions did not alter activity (62,63). N145 is the residue in our model that appears to be able to contact the unpaired thymine base (the partner
of the flipped out adenine). In the second TRD, the equivalent loop is T354SGQKGISGK363, and Q357 is able to contact the unpaired thymine in the complementary strand for the second part of the target sequence, 5’-GATC-3’.

HsdM to DNA contacts. Wilcock et al. (66) substituted amino acids within the well-characterized, conserved MTase motifs within the catalytic domain of HsdM. The substitution in the AdoMet binding motif I, G177D, affected cofactor binding and the model shows motif I in close contact with the AdoMet (Figure 3f). Changes at locations N266 or F269 in the catalytic motif IV led to a loss of activity. In the model of M.EcoKI with DNA bound, these locations can be seen adjacent to the flipped-out adenine (Figure 3f). In addition, F345 is close to the flipped-out adenine. A similar sandwich of a flipped-out adenine between two aromatic side chains was observed in the crystal structure of the M.TaqI MTase (67).

Analysis of the M.EcoKI–ocr model

This atomic model (Figure 1) is more open than the model with DNA (Figure 2) as there are fewer contacts between the three subunits of M.EcoKI. In particular, the interface between the two HsdM is formed only by helix 64–78, and the C-terminal helices of HsdM, 487–507, appear to have minimal contact with the coiled-coil region of HsdS. It appears that ocr is slightly too large to be as easily accommodated as the DNA molecule. As the subunit interfaces are smaller, one might say that the MTase is almost ‘bursting at its seams’ as it attempts to enwrap the ocr protein (47). This suggests that a clamp-like open and shut mechanism is used by M.EcoKI to bind DNA or ocr.

Mechanism for binding DNA or ocr

The EM structure makes it clear that M.EcoKI can encircle DNA and, to a lesser extent, ocr. Therefore, the structure must open up to allow the DNA (or the ocr protein) access to the TRDs on HsdS and the catalytic site on HsdM. Simple normal mode calculations were performed on HsdS to examine the potential role of the coiled-coil region in allowing such an opening (data not shown). These calculations showed that the lowest mode was a flexing of the coiled-coil moving the TRDs closer together and then further apart. The next lowest mode was a twisting of the coiled-coil around its axis. Noting that the interfacial DNA is approximately modelled by Obarska et al. (37) to produce a gap of approximately 0.7 nm in our model, this opening and closing motion, which must open a gap at least 2 nm wide between HsdM subunits to allow double-stranded DNA to pass through, is facilitated by twisting and bending of the coiled-coil region in the HsdS subunit.

Comparison to the Type IC enzyme EcoR124I

Kneale (17) originally recognized the 2-fold symmetry imposed upon Type I enzymes via the structure of the S subunit. This symmetry defines the location of the M subunits and was exploited by Obarska et al. (19) to produce a rather opened-out model of the MTase where the N-terminal domains of each HsdM subunit were speculatively placed some distance from both the DNA and any contact with the other HsdM subunit. Although based upon an erroneous structure of an HsdM subunit (pdb 2ar0) is incomplete and partially incorrectly assigned, see Supplementary Data), the model could, if correct, explain the large compaction of M.EcoR124I observed by small angle neutron scattering when DNA is added to the enzyme (67). This compaction was suggested to be the movement of the HsdM subunits swinging in towards the DNA and closing like a clamp to trap the DNA rather like the closed state we have observed in our structure (Figure 1). Thus, it is probable that the structure of Type I MTases fluctuates between an open form as approximately modelled by Obarska et al. (19) and the closed form observed here and by Taylor et al. (67). Normal mode analyses of our model make it clear that this opening and closing motion, which must open a gap at least 2 nm wide between HsdM subunits to allow double-stranded DNA to pass through, is facilitated by twisting and bending of the coiled-coil region in the HsdS subunit.

More recently, small-angle neutron scattering data and computer modelling of the HsdR subunit of EcoR124I have been presented (20). The HsdR subunit appears to be a disc-shaped protein but a unique atomic model could not be constructed. Presumably, one of these subunits could be placed on either side of the MTase structure so that the ends of the DNA helix would contact the helicase domains in the HsdR subunit. Until a unique structure of the HsdR subunit is presented and more data are acquired on its contacts to the MTase, further speculation on the complete structure of the Type I R/M enzymes seems premature. However, we note that we are currently making progress towards this goal.

CONCLUSION

The new EM structure and computational model of M.EcoKI rationalize, for the first time, a large body of experimental data obtained using many different methods over many years. A mechanistic explanation of the Type I MTase enzymes is suggested by the model, which clearly indicates locations for further analyses such as the HsdM–HsdS and HsdM–HsdM interfaces. When combined with data on M.EcoR124I (67), the model also suggests that a dynamic opening and closing of the protein, driven by a flexing and twisting of the conserved coiled-coil region within HsdS, is required to open up the HsdM–HsdM interface to allow either DNA binding or attack of the ocr antirestriction protein, a protein that ‘disguises itself’ as DNA. The model also provides a basis for completing or constructing models of other Type I MTases from other
families such as the Type IC enzyme M.EcoR124I (19) and the Type IB enzyme M.EcoAI.

Data depositions

Atomic coordinates from M.EcoKI–ocr and M.EcoKI–DNA complexes are available from the authors or ftp://ftp.genesilico.pl/iamb/models/MTases/EcoKI/. The EM map is deposited with the EMDB (www.ebi.ac.uk/msd) with code 1543.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was initiated at the Isaac Newton Institute for Mathematical Sciences Workshop on ‘Statistical Mechanics of Molecular and Cellular Biological Systems’, January to July 2004. We thank Professor Noreen Murray and numerous colleagues in her laboratory for countless discussions on the structure of EcoKI.

FUNDING

We gratefully acknowledge support from the Biotechnology and Biological Sciences Research Council (BB/D001870/1 to D.T.F.D. and J.T.), the Wellcome Trust (GR080463MA to D.T.F.D.) and the Royal Society (BB/D001870/1 to D.T.F.D. and J.T.). We gratefully acknowledge support from the Biotechnology and Biological Sciences Research Council for Mathematical Sciences Workshop on ‘Statistical Dependence Methyltransferases: Structures and Functions.’

REFERENCES

5. Murray,N.E. (2000) Type I restriction systems: sophisticated models such as the Type IC enzyme M.EcoR124I (19) and the Type IB enzyme M.EcoAI.
the EcoKI endonuclease visualized by atomic force microscopy. 


Funct. Integr. Genomics, 6, 186–201.


Curr. Opin. Microbiol., 8, 466–472.


Ultramicroscopy, 21, 111–123.


