

# Chapter 6

## Structural characterisation

### 6.1 Introduction

Small angle neutron scattering (SANS) was carried out on both the individual components of the MTase and ENase, and on the multisubunit complexes in the presence and absence of a 30 bp DNA duplex, in order to elucidate more fully the structural change that takes place upon DNA binding. Finally, *ab initio* shape determination by dummy atom modelling allowed us to discover the change in subunit positioning upon DNA binding.

Predictive structural modelling of the S subunits of EcoR124I<sub>NT</sub> and EcoR124I based on published crystal structures of a homologous S subunit were also undertaken. In addition hydrodynamic models of a number of crystal structures from Type I S subunits were carried out to calculate parameters that could be related to ones obtained for EcoR124I<sub>NT</sub>.

Finally, attempts to crystallise S3 and S3/M in the presence and absence of a 30 bp DNA duplex are also discussed.

### 6.2 Small Angle Neutron Scattering (SANS)

#### 6.2.1 Introduction

SANS experiments of the free subunits and complexes of the EcoR124I<sub>NT</sub> R-M system in the presence and absence of DNA were undertaken in collaboration with the Deuteration Laboratory of the Large Scale Structures Group at the Institut Laue-Langevin (ILL), Grenoble, France, using the D22 instrument.

Small angle X-ray scattering (SAXS) has previously revealed a 60 Å decrease in the overall dimensions of the wild-type M.EcoR124I MTase upon DNA binding (Taylor *et al.*, 1994). Therefore, SANS experiments were carried out in order to

determine the change in the positioning of the S3, M and R subunits as complexes of S3/M and S3/M/R, upon binding to DNA.

After the measurement of the scattering data, three parameters can be determined: the radius of gyration,  $R_g$ ,  $D_{max}$ , the longest dimension of a particle and the pair distance distribution function,  $p(r)$ , which describes the distribution of all inter-atomic distances in the particle. Together with specific deuteration and contrast variation, these structural parameters can also be obtained for individual subunits within the complexes.

### 6.2.2 Sample preparation and experimental design

In all experiments discussed in section 6.2, proteins and protein-DNA complexes were purified and formed as described in section 2.10. Table 6.1 lists a summary of the experiments carried out.

### 6.2.3 Data correction and analysis

The data were first corrected for sample transmission. Background scattering and scattering from the empty cell were subtracted from the radially-averaged sample data. These values were then divided by the scattering due to water and finally the scattering due to the cell containing only buffer was subtracted. The data was also normalised for concentration. At low angle, the isotropic scattering data can be expressed as the Guinier approximation:

$$I(s) = I(0) \exp -1/3 R_g^2 s^2 \quad (6.1)$$

Where  $s$ , is the scattering vector ( $4\pi\sin\theta/\lambda$ ). The isotropic scattering intensity,  $I(s)$  or  $I(Q)$ , can be transformed to give the pair distance distribution function,  $p(r)$ . The radius of gyration is defined as the root-mean-square-distance of an array of atoms or groups from their common centre of gravity. By plotting a straight line through a Guinier plot,  $\ln I(s)$  versus  $s^2$ , the slope gives the radius of gyration. The Guinier region is the  $s$  (or  $Q$ ) range in which the plot is valid. Signs of aggregation can also be detected from this plot.

Fully hydrogenated samples	Protein conc. (mg mL <sup>-1</sup> )	Pathlength (mm)	% D <sub>2</sub> O
S3	4.7	2	100
S3	0.6	2	100
R	1.36	2	100
M	4.22	2	100
S3/M/R	1.05	2	100
S3/M plus DNA	4.20	2	100
S3/M/R plus DNA	1.03	2	100
S3/M	2.11	1	100
S3/M/R	1.05	1	100
S3/M plus DNA	2.09	1	100
S3/M/R plus DNA	1.03	1	100

Samples with deuterated S3	Protein conc. (mg mL <sup>-1</sup> )	Pathlength (mm)	% D <sub>2</sub> O
dS3/M	1.44	2	100
dS3/M	0.72	1	100
dS3/MR	1.05	1	100
dS3/M plus DNA	2.09	1	100
dS3/M/R plus DNA	1.03	1	100
dS3/M	2.11	1	40
dS3/M/R	1.05	1	40
dS3/M plus DNA	2.09	1	40
dS3/M/R plus DNA	1.03	1	40
<b>Buffers</b>			
IEX1	N/A	2	100
IEX1	N/A	1	100
IEX1	N/A	1	40

**Table 6.1: Summary of SANS experiments carried out.** All measurements were carried out at both 2 m and 8 m, in either 1 or 2 mm pathlength cells, total volume either 400  $\mu$ L or 200  $\mu$ L, respectively.

#### 6.2.4 Contrast variation

Contrast variation is accomplished by varying the percentage of H<sub>2</sub>O/D<sub>2</sub>O in the buffer. Specific deuteration can also be used to highlight specific fragments or domains of a protein. In the following experiments, it is possible to ‘contrast-match out’ the deuterated S3 subunit if it is in 100% D<sub>2</sub>O, and thus only measure the scattering from the M subunit within a complex of dS3pM. Alternatively, it is possible to measure the scattering due to S3 within a complex of S3/M in 40% D<sub>2</sub>O, due to the protonated M subunit being ‘matched out’. Table 6.2 shows the scattering length densities of various components.

#### 6.3.5 SANS results

Figures 6.1 and 6.2 show the scattering curves measured for the various complexes. Figures 6.3 and 6.4 show the corresponding  $p(r)$  functions, from which  $D_{\max}$  [when  $p(r) = 0$ ] can be calculated. Table 6.3 summarises the radii of gyration ( $R_g$ ) and  $D_{\max}$  for the data collected. The data obtained for measurements carried out in a 100% deuterated buffer for the fully protonated S3/M complex and the selectively deuterated dS3pM gave  $R_g$  values of 56.3 Å and 47.8 Å, respectively. The value for the fully protonated complex is in very close agreement with the value of 56 Å obtained for the wild-type MTase EcoR124I by SAXS (Taylor *et al.*, 1994).

A noticeable decrease in the radius of gyration is seen with the addition of R to the fully protonated pS3/M complex in 100 % D<sub>2</sub>O, from 56.3 Å to 42.0 Å, and a further decrease to 37.4 Å is observed upon the addition of DNA. It appears that there are rearrangements of the subunits of the S3/M complex upon addition of R to form the ENase prior to binding DNA. This could be due to further compaction of the S3/M ‘core’ through R stabilising the interaction between the S3 and M subunits.

For the case of S3/M, an overall decrease in  $D_{\max}$  of 50 Å (from 195 Å to 145 Å) is seen on binding DNA. This is comparable to the wild-type enzyme EcoR124I, where there is a decrease from 180 Å to 112 Å, indicating that a similar structural re-arrangement is taking place.

<b>Component</b>	<b><math>\rho</math> in H<sub>2</sub>O (10<sup>10</sup> cm<sup>-2</sup>)</b>	<b><math>\rho</math> in D<sub>2</sub>O (10<sup>10</sup> cm<sup>-2</sup>)</b>	<b>Match point in % D<sub>2</sub>O</b>
<b>H<sub>2</sub>O</b>	0.6	-	-
<b>D<sub>2</sub>O</b>	6.4	-	-
<b>50% sucrose</b>	1.2	-	-
<b>Lipids</b>	0.3	6.0	10 – 15%
<b>Proteins</b>	1.8	3.1	40%
<b>Nucleic acids</b>	3.7	4.8	70%

**Table 6.2: Neutron scattering length densities of various biological components.**

A similar set of observations is seen when the same measurements are carried out with complexes containing deuterated S3 in 100% D<sub>2</sub>O. Upon DNA binding, the radius of gyration of dS3/M decreases from 47.8 Å to 32.1 Å. The smaller radii of gyration in both cases are to be expected due to the change in the ‘contrast-matched out’ S3 subunit, *i.e.* only scattering from the M subunit ± DNA is recorded. This implies a major compaction of the M subunits when the MTase binds to DNA.

No difference in either the radius of gyration or  $D_{\max}$  is seen with dS3/M in the presence or absence of DNA in 40% D<sub>2</sub>O, when the M subunits are contrast matched out. This suggests that the S3 subunits do not change shape upon DNA binding and that the large compaction of the MTase following DNA binding is due only to movement of the M subunits.

#### 6.2.6 *Ab initio* modelling

In order to create a model of S3/M from the neutron scattering curves, the program DAMMIN (Svergun, 1999) was used. In this approach, a sphere of diameter equal to the maximum particle size is filled with M densely packed small spheres or dummy atoms, where  $M \approx (D_{\max}/r_0)^3 \approx 10^3$ . DAMMIN tries to find a compact dummy atom configuration that is compatible with the scattering curve, with a penalty for loose or disconnected configurations.

The preliminary results from the *ab initio* modelling of S3/M with and without the 30 bp DNA duplex containing the recognition sequence for S3/M are shown in figure 6.5. Shown in green is the fully protonated S3/M complex in 100 D<sub>2</sub>O, which represents the whole complex. Shown in blue is the perdeuterated complex in 100% D<sub>2</sub>O, where the S3 subunit was selectively deuterated. Thus in this case only the M subunit is “seen”. The crystal structure for the S subunit of *Methanococcus jannaschii* (Kim *et al.*, 2005) was superimposed onto the S3/M model and to some extent fits well with the position of the S3 subunits. Upon DNA binding, there is a decrease in the  $D_{\max}$  from 200 Å to 150 Å. The model again illustrates the very large compaction of the S3/M complex upon DNA binding.

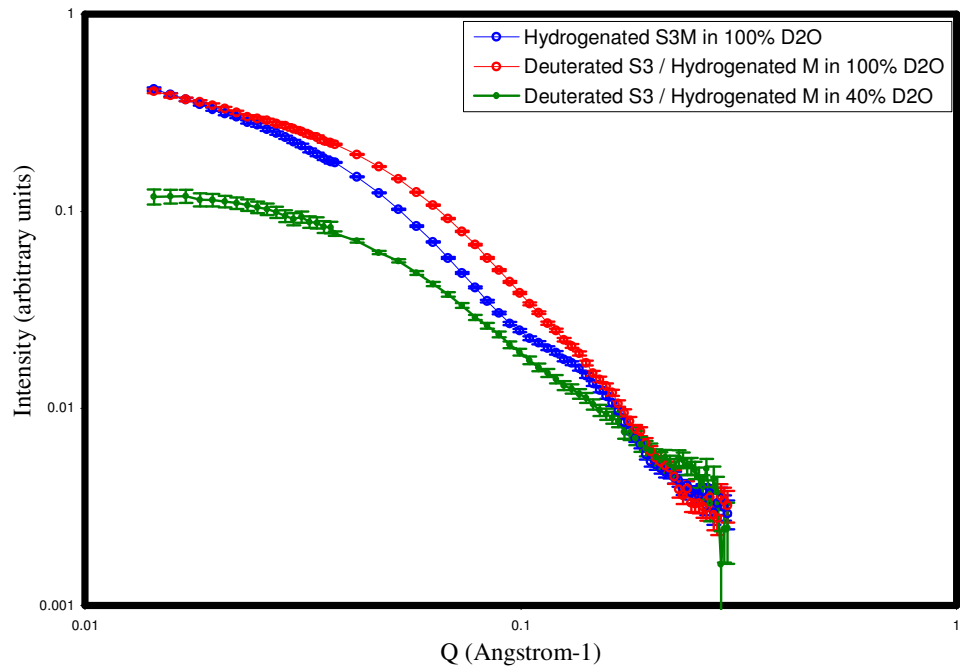
Protonated samples	% D <sub>2</sub> O	R <sub>g</sub> (Å)		D <sub>max</sub> (Å)	
		-DNA	+ DNA	- DNA	+ DNA
S3/M	100	56.3 ± 1.2	49.1 ± 0.4	195	145
S3/M/R	100	42.0 ± 1.7	37.4 ± 1.0	140	140

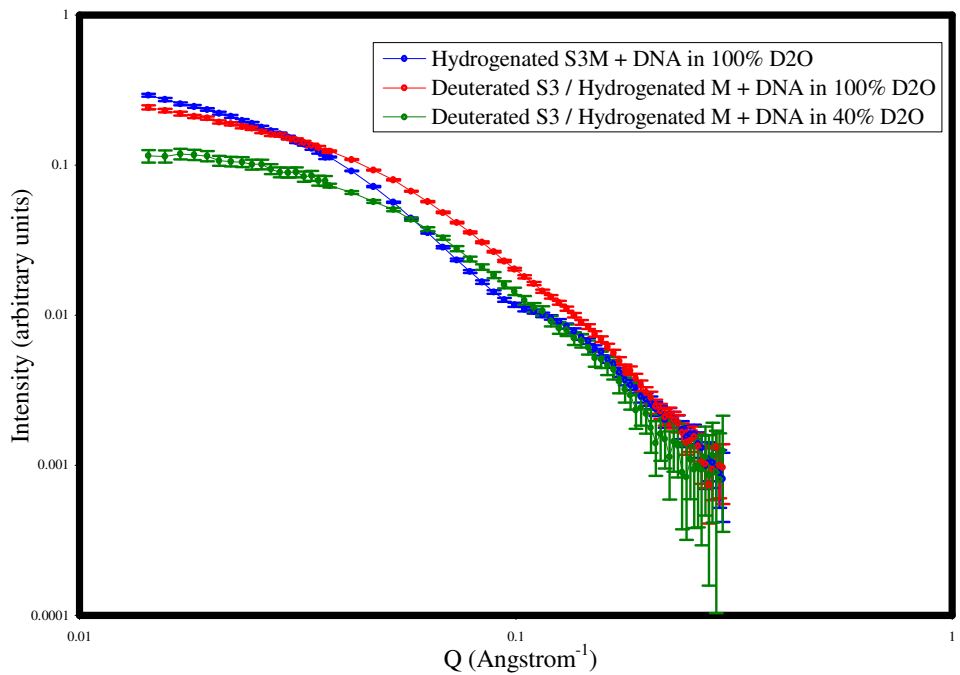
Deuterated samples	% D <sub>2</sub> O	R <sub>g</sub> (Å)		D <sub>max</sub> (Å)	
		- DNA	+ DNA	- DNA	+ DNA
dS3/M	100	47.8 ± 1.9	32.1 ± 0.4	195	95
dS3/M/R	100	37.8 ± 1.1	37.4 ± 1.0	140	140
dS3/M	40	31.0 ± 1.1	33.3 ± 1.3	110	110

**Table 6.3: Summary of the experimental radius of gyration (Å) and D<sub>max</sub> (Å) values.** Experiments were performed in the presence (+) and absence (-) of the DNA recognition sequence for EcoR124I<sub>NT</sub>.

(A)

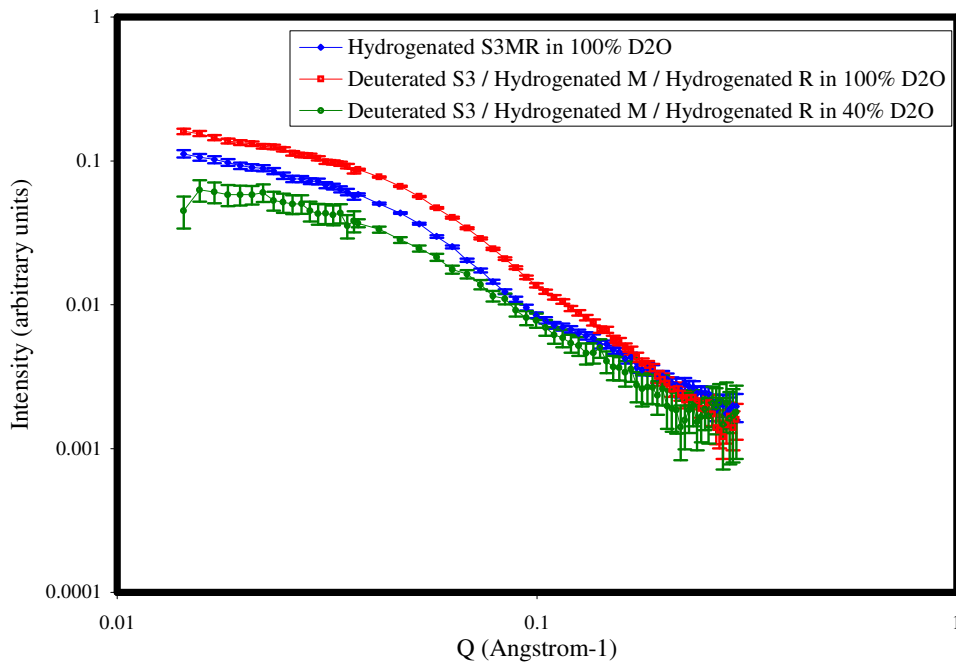


(B)

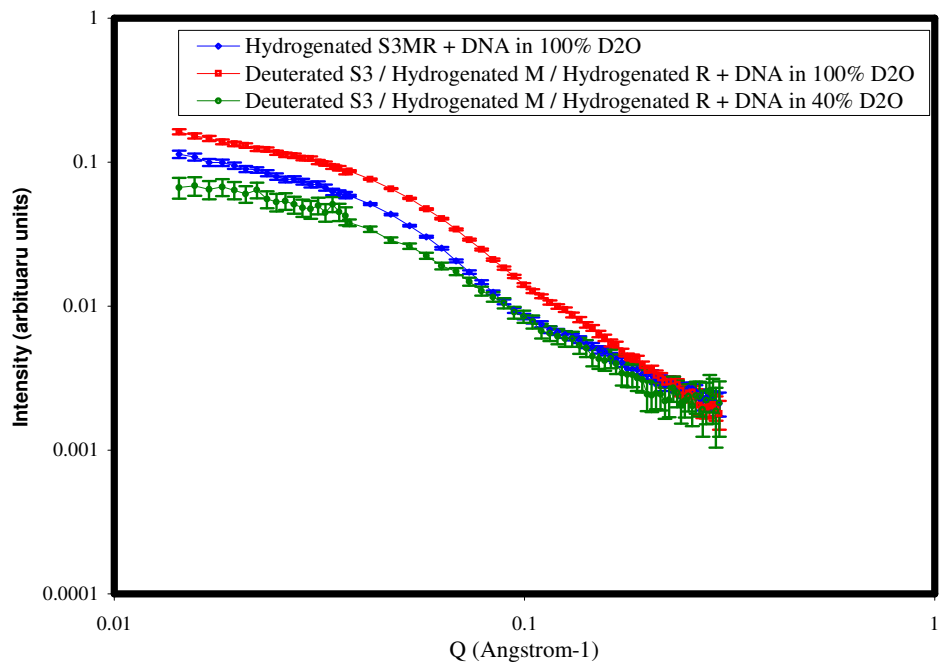


**Figure 6.1: SANS scattering curves for the fully protonated and perdeuterated S3/M complex in the presence and absence of DNA. (A) with DNA and (B) without DNA. Both were measured in 40% and 100% D<sub>2</sub>O.**

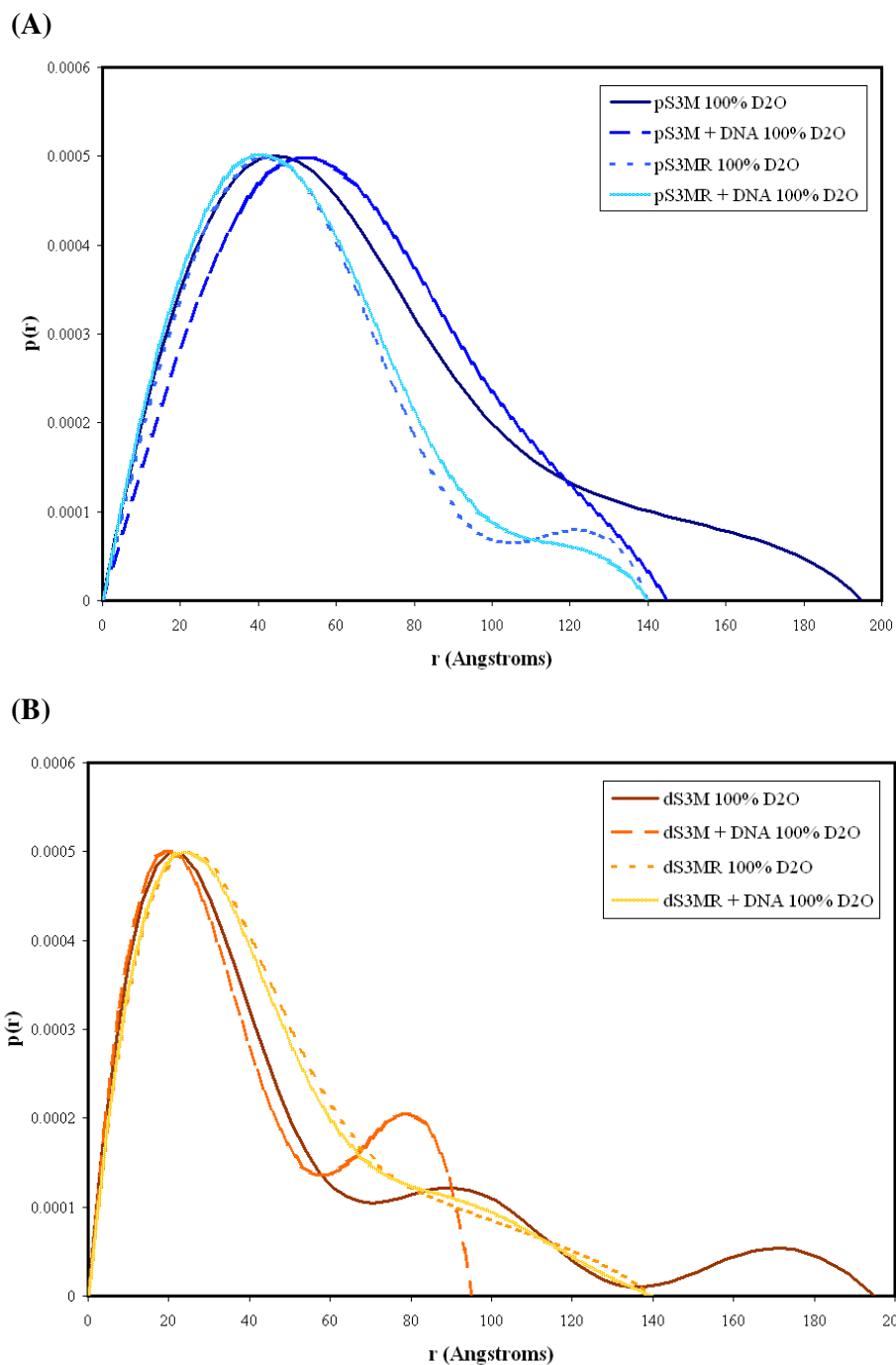
(A)



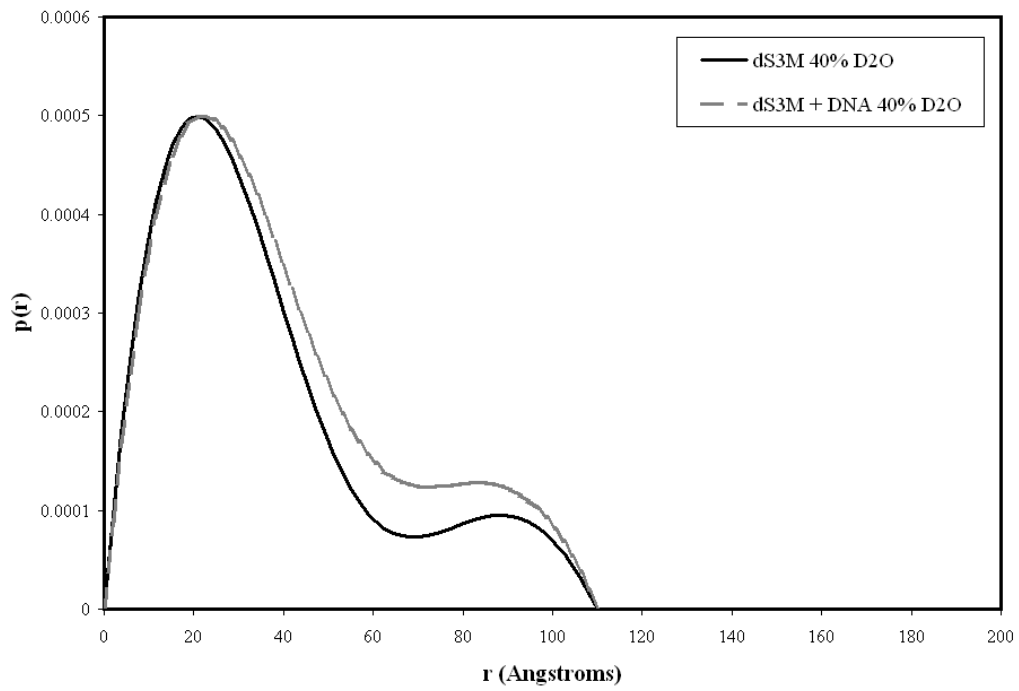
(B)



**Figure 6.2: SANS scattering curves for the fully protonated and perdeuterated S3/M/R complex in the presence and absence of DNA. (A) with DNA and (B) without DNA. Both were measured in 40% and 100% D<sub>2</sub>O.**



**Figure 6.3: Distance distribution function of (A) the fully protonated S3/M and S3/M/R complexes in 100% D<sub>2</sub>O and (B) perdeuterated S3/M. Both are shown in the presence and absence of a 30mer DNA duplex containing the recognition sequence for S3/M at a 1:1 ratio.**



**Figure 6.4: Distance distribution function of perdeuterated S3/M in 40% D<sub>2</sub>O in the presence and absence of a 30mer DNA duplex containing the recognition sequence for S3/M at a 1:1 ratio.**

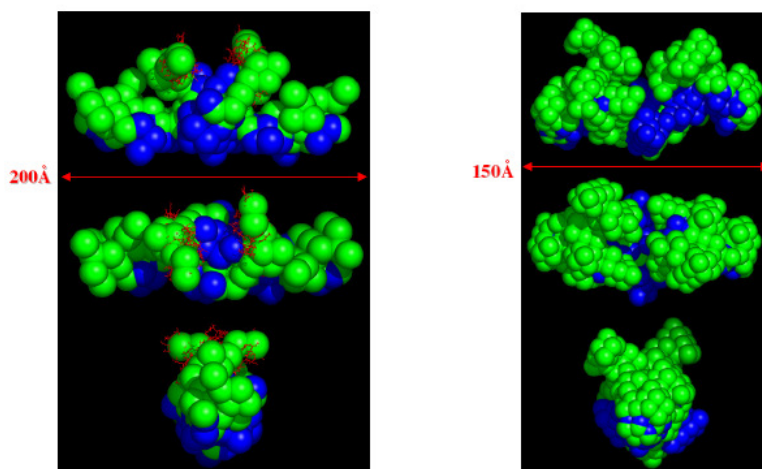


Figure 6.5: Low resolution structure of S3/M in the presence and absence of DNA. The models were restored from the scattering pattern using the program DAMMIN (Svergun, 1999). In green is fully hydrogenated S3M in 100% D<sub>2</sub>O, in blue is deuterated S3/hydrogenated M in 100% D<sub>2</sub>O and in red is S subunit crystal structure of *Methanococcus jannaschii* (Kim *et al.*, 2005).

## 6.3 Subunit homology and model prediction

### 6.3.1 S subunit

The programme JPred2 (Cuff *et al.*, 1998) was used to predict the secondary structure of the S subunit of EcoR124I (figure 6.6), which was found to be in close agreement with the secondary structure of the solved S subunit structure of *Methanococcus jannaschii* (Kim *et al.*, 2005). Furthermore, a global pair-wise alignment shows there is some homology between the amino acid sequence of the S subunit from EcoR124I and that of the S subunit from *M. jannaschii*, with an identity of 23% and a similarity of 40% (figure 6.7).

Therefore a homology-based approach using the program SwissModel (Schwede *et al.*, 2003) was used to predict the structure of the S subunits of EcoR124I, EcoR124I<sub>NT</sub> and AhdI based on the crystal structure of the S subunit of *M. jannaschii*. However it failed to fully model these proteins (data not shown) although it did successfully predict their coiled-coil regions.

However a complete model of the S subunit of EcoR124I (figure 6.8A) was obtained using the program GeneSilico (Kurowski *et al.*, 2003) and showed a similar structural arrangement to the solved S subunit structure of *M. jannaschii*, which incidentally shows flexibility when the two subunits in the asymmetric unit (chains A and B) are overlaid (figure 6.8B).

The web-based program QuickPhyre (Protein Homology/analogy Recognition Engine - Kelley *et al.*, 2000) was used to find structurally related proteins. The coiled-coil structure was found to occur in many other proteins, for example the *Pyrococcus furiosus* Rad50 ATPase domain (figure 6.8C).

Figure 6.9 shows a hydrodynamic model created using the program HYDROPRO (Garcia de la Torre *et al.*, 2000) based on the crystal structure of the S subunit of *M. jannaschii*. This allows the determination of theoretical values for the radius of gyration and the sedimentation coefficient. These values can then be compared to the experimental  $S_{20,w}$  (obtained from



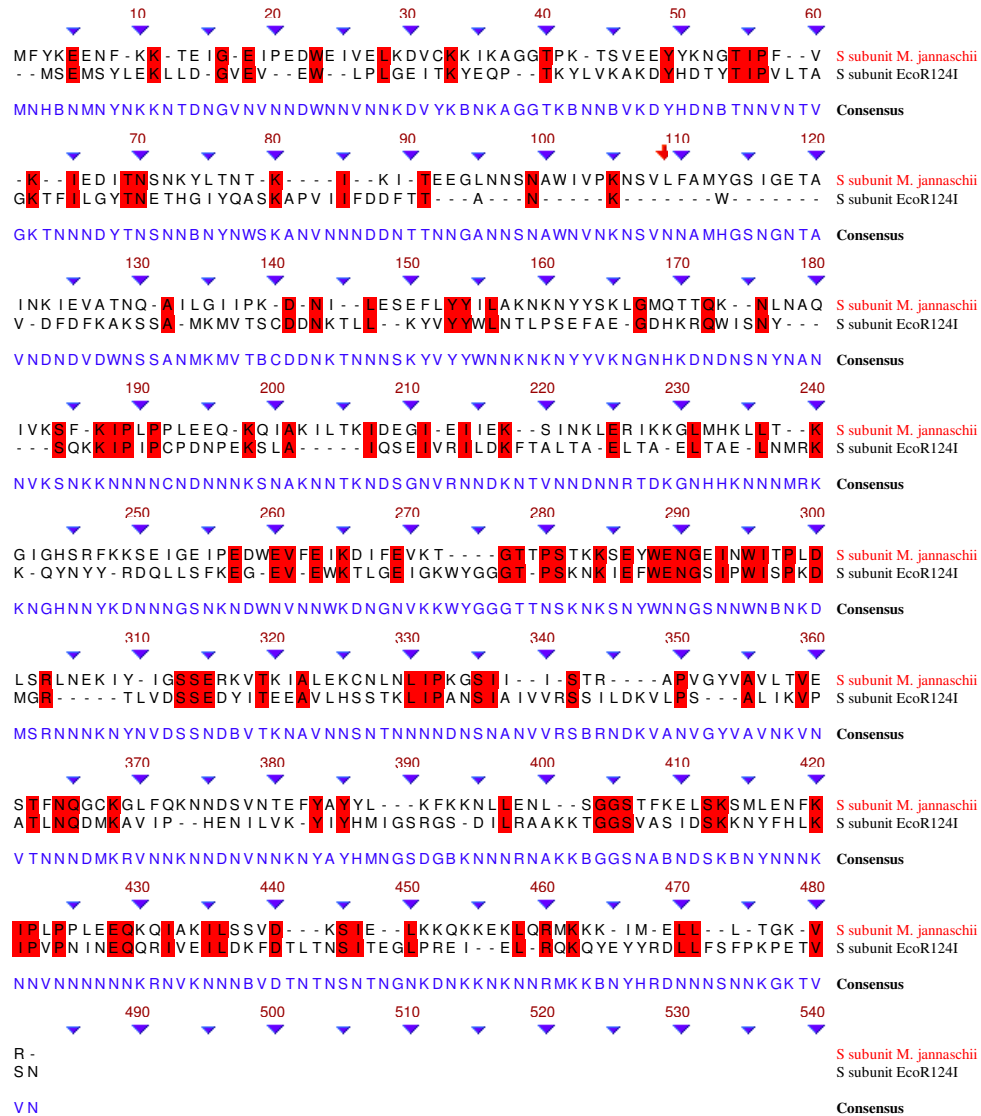
analytical ultracentrifugation) and  $R_g$  (obtained from SANS) for other systems, such as EcoR124I<sub>NT</sub>.

The model is initially built from the atomic coordinates, by replacing non-hydrogen atoms with spherical beads of fixed radius. The particle that is formed consists of overlapping spheres, which is then reproduced as a shell model. A value of 32.3 Å was obtained for the radius of gyration for the crystal structure of the S subunit of *M. jannaschii*. This corroborates the SANS results obtained for dS3pM in 40% D<sub>2</sub>O, where the protonated M is ‘contrast-matched out’ and this  $R_g$  value of 31 Å represents S3. The sedimentation coefficient of 3.3 S obtained by this method is also very similar to the experimental sedimentation coefficient of 3.42 S measured for S3 (Smith *et al.*, 2001).

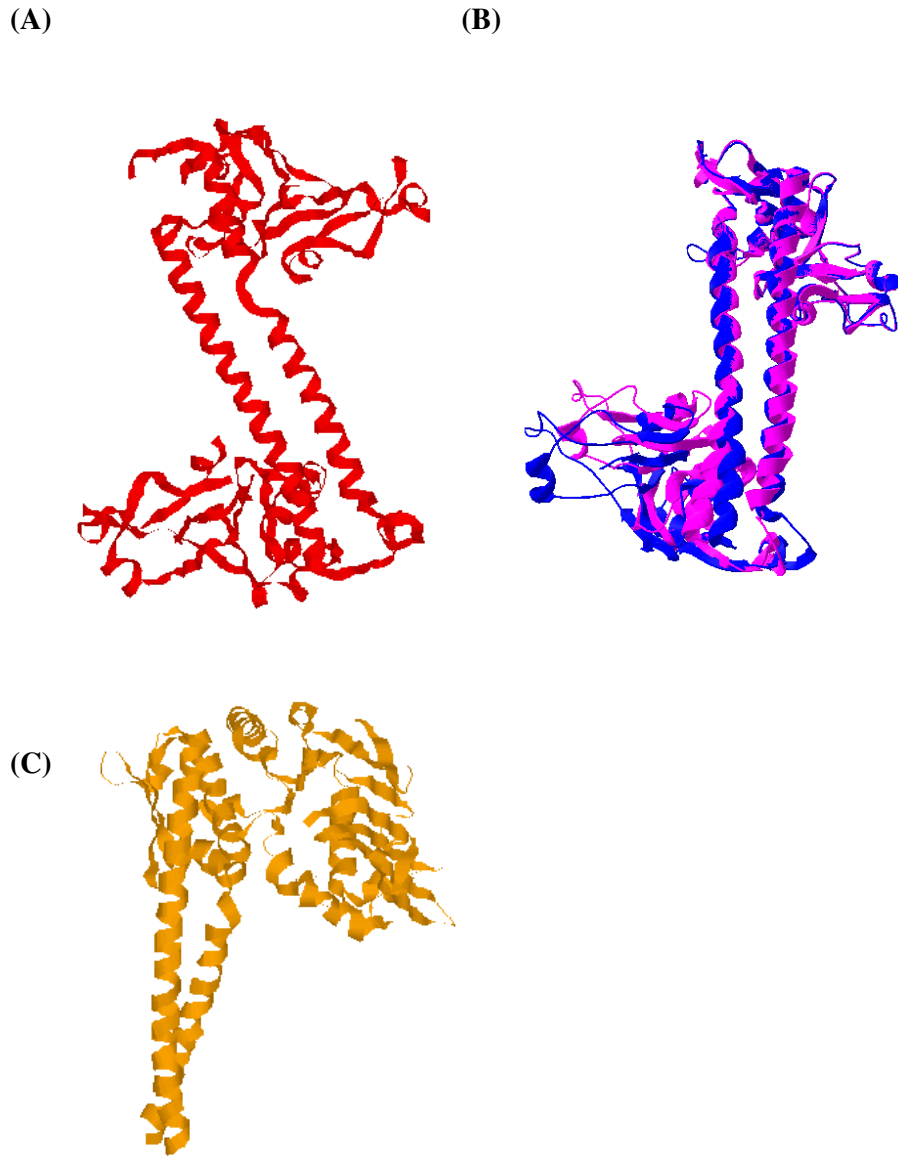
This suggests, not only that the S3 dimer closely resembles the structure of the S subunit of *M. jannaschii*, but that the structure of the S3 dimer does not change when forming the MTase.

### 6.3.2 The M subunit

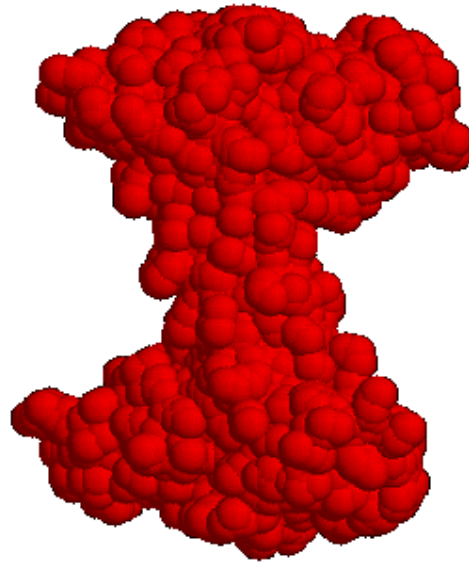
The recently reported M subunit structure, pdb entry 2AR0 (Rajashankar *et al.*, *to be published*), has been solved to a resolution of 2.8 Å (figure 6.11). Like the M subunit of EcoR124I<sub>NT</sub> (this thesis), it forms a dimer in solution. The dimer is mainly alpha helical, comprising 20 alpha helices but also twelve sheet regions. A global pair-wise alignment of the M subunits of EcoR124I and EcoKI show that they share 20 % identity and 32 % similarity (figure 6.10). Based on the same approach for the S3 subunit, a hydrodynamic model of the crystal structure of the M subunit from EcoKI was constructed (figure 6.12). Values of 27 Å for the radius of gyration, 35 Å for the hydrodynamic radius and a  $S_{20,w}$  of 4.04 S for a monomer of M were obtained by this method. However the program was unable to calculate these parameters for the dimer, and so cannot be compared with the experimental SANS and SV parameters.



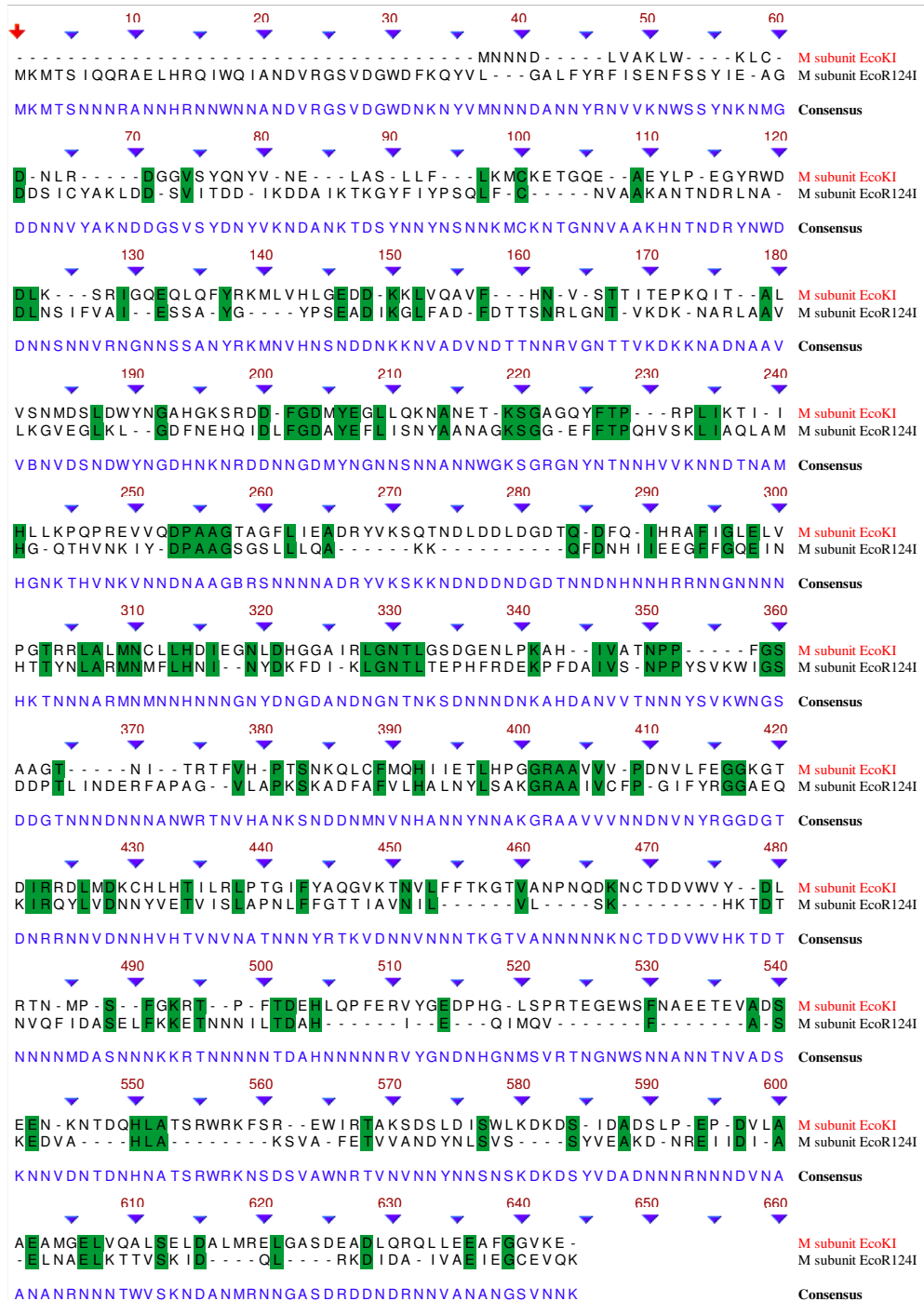
**Figure 6.7: A global pair-wise alignment of the S subunits of *Methanococcus jannaschii* and EcoR124I.** Below in blue is the consensus sequence. In red, the conserved amino acid residues.



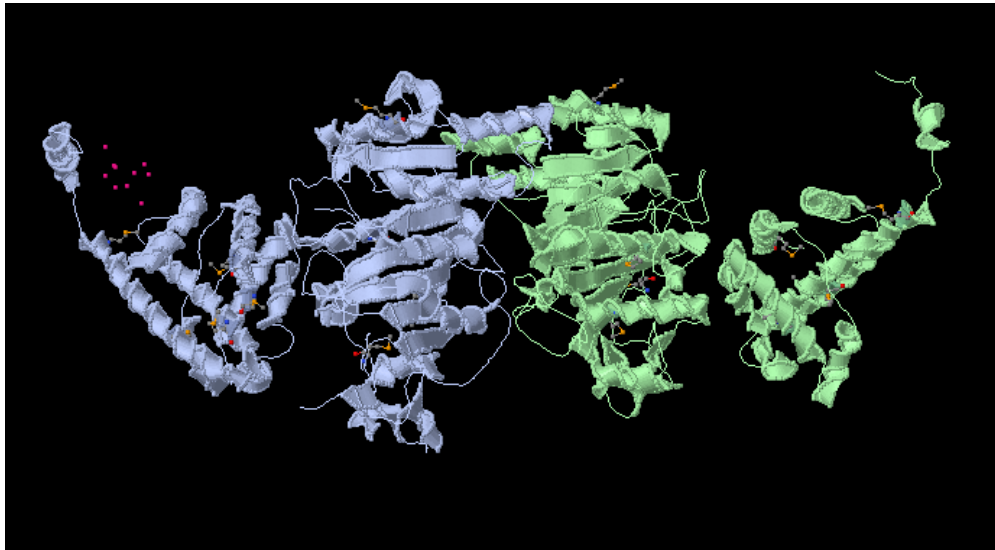
**Figure 6.8:** (A) **Predicted structure of the S subunit of EcoR124I.** The models were obtained using GeneSilico, by comparison with the published structure of the S subunit of *M. jannaschii*. (B) **An overlay of each of monomer of the crystal structure of *M. jannaschii*.** A ribbon diagram represents the tertiary structure. In blue, chain A, and in pink, chain B. (C) **Crystal structure of the *P. furiosus* Rad50 ATPase domain (PDB entry 1II8).** The structure was found using the @TOME meta-server.



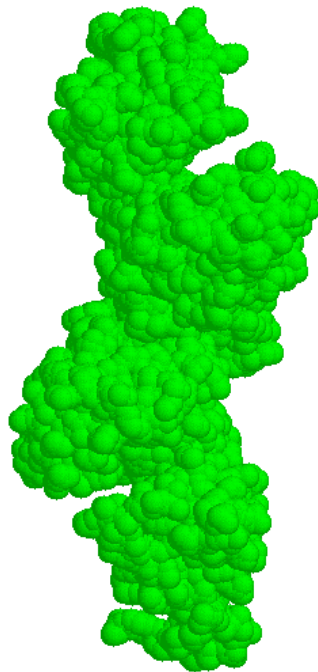
**Figure 6.9: Hydrodynamic model of *Methanococcus jannaschii* (chain A, PDB 1YF2).** The hydropro program (Garcia de la Torre *et al.*, 2000) was used to calculate a primary hydrodynamic model using the published crystal structure of *M. jannaschii* by the shell-model technique. Values for  $R_g$  (32 Å),  $R_h$  (3.5 Å) and  $S_{20,w}$  (3.3 S) were obtained by this method.



**Figure 6.10: A global pair-wise amino acid alignment of the M subunits of EcoKI and EcoR124I.** Below in blue is the consensus sequence. In green, the conserved amino acid residues.



**Figure 6.11: Crystal structure of the M subunit of the Type IA restriction enzyme EcoKI.** (Rajashankar *et al.*, *to be published*).



**Figure 6.12: Hydrodynamic model of a dimer of M of EcoKI (PDB 2AR0).** The hydropro program (Garcia de la Torre *et al.*, 2000) was used to calculate a primary hydrodynamic model using the crystal structure of the M subunit of EcoKI by the shell-model technique.

## 6.4 Crystal trials

Previous attempts to crystallise S3 and S3/M using Hampton Crystal Screens 1 and 2 only produced precipitation. This screen is based on the sparse matrix, which varies salt, pH and precipitant (Lindwall *et al.*, 2000). Based on the concentrations used it was decided to use a lower concentration of protein and instead try Structure screens 1 and 2 and Pact screens 1 and 2 (Molecular Dimensions) again by the hanging drop vapour diffusion method. However no suitable crystals were obtained for S3 or S3/M with or without the 30 bp EcoR124I<sub>NT</sub> DNA sequence at concentrations up to 5 mg/mL.

## 6.5 Discussion

From the SANS data we have collected three parameters from the analysed scattering data, the radius of gyration ( $R_g$ ), the overall length ( $D_{max}$ ) and the distribution function  $p(r)$  for the complexes of the EcoR124I<sub>NT</sub> R-M system in the presence and absence of DNA. Furthermore, by *ab initio* modelling to the SANS data, the change in the positioning of the subunits of the relevant complexes upon DNA binding was also investigated. By creating hydrodynamic models of recently solved S and M subunits of type I R-M systems, parameters such as the sedimentation coefficient ( $s_{20,w}$ ), radius of gyration ( $R_g$ ) and hydrodynamic radius ( $R_h$ ) could then be compared to experimental values obtained by SANS and SV experiments for the EcoR124I<sub>NT</sub> R-M system.

### 6.5.1 MTase

A  $R_g$  of 56 Å was obtained for S3/M, which is identical to that obtained for the wild-type MTase M.EcoR124I by SAXS indicates both are adopting the same conformational arrangement, even though there is no covalent link between the TRDs in the case of S3.

### 6.5.2 ENase

Upon addition of the R subunit to S3/M a decrease from 56 Å to 42 Å occurred, suggesting further rearrangement of the complex prior to DNA binding. The M subunits appear to be more compact when R is bound, whether or not DNA is

present. To investigate this further, and separately observe the  $R_g$ 's of the M, S3 and R subunits, would require SANS experiments with additional combinations of deuterated/protonated subunits (*e.g.* addition of deuterated R to protonated S3/M in 100 %  $D_2O$ ).

### 6.5.3 DNA binding

When measurements are carried out with S3/M in the presence of its recognition sequence, the overall dimensions of the enzyme ( $D_{max}$ ) decrease from 200 Å to 150 Å. This is comparable to the change measured by small angle X-ray scattering of M.EcoR124I, in which there was a decrease from 180 Å to 112 Å (Taylor *et al.*, 1994). However it is notable that we do not see any change in  $R_g$  of the S3 subunit in 40 %  $D_2O$ , when we are effectively measuring scattering from S3, in the presence and absence of DNA. This indicates that once the M subunit is bound to the S3 subunit, S3 maintained in a rigid conformation through the stabilising interactions with the M subunit. The results imply that the large decrease in  $R_g$  we observe upon DNA binding of the MTase is solely due to a structural change in the M subunits.

### 6.5.4 Hydrodynamic modelling

Both the crystal structure and hydrodynamic models of the S subunit of *M. jannaschii* (Kim *et al.*, 2005) and the M subunit of EcoKI (Rajashankar *et al.*, *to be published*) resemble the final models of S3 and M subunit of EcoR124I<sub>NT</sub> obtained by SANS. Indeed the parameters obtained for the S subunit of *M. jannaschii* are in very close agreement with those measured for S3 ( $s_{20,w} = 3.3$  S and 3.42 S respectively and  $R_g = 32.3$  Å and 31 Å respectively).