2 Materials and Methods

2.1 Materials

2.1.1 Bacterial strain

Strain: *E. coli* SCS1 (*hsdR hsdM*$^+$ *recA1*, Stratagene)

2.1.2 Expressionvector

Vector: pMS119EH (pMB1 replicon, p_{tac}/*lacI*, Ap^r)

2.1.3 Medium and antibiotic resistance

LB Medium: bacto-tryptone 10 g, bacto-yeast extract 5 g, NaCl 10 g, adjust pH to 7.0, add H_2O to 1 liter, sterilize by autoclaving for 20 min.

YT Medium: bacto-tryptone 8 g, bacto-yeast extract 5 g, NaCl 2.5 g, adjust pH to 7.0, add H_2O to 1 liter, sterilize by autoclaving for 20 min.

Antibiotic: Ampicillin (100mg/l)

2.1.4 Reagents and buffers

All chemicals used in this study were of p. a. quality and solutions were prepared with Milli-Q water. Buffer A used for ATPase activity assays and nucleotide binding tests at pH 5.8 contained 40 mM MES/NaOH, 10 mM $MgCl_2$, 60 mM NaCl. Buffer B used for ATPase activity assays and nucleotide binding tests at pH 7.6 contained 40 mM Tris/HCl, 10 mM $MgCl_2$. Buffer C used for helicase activity measurements contained 40 mM MES/NaOH, pH 5.6, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 50 µg/ml bovine serum albumin, 0.02% (wt/wt) Brij-58.

2.1.5 Chemicals and enzymes

Ammonium molybdate tetrahydrate was purchased from Fluka Biochemika and malachite green from Sigma. ATPγS, ATP, ADP and AMPPNP were purchased from Sigma and [γ-^{32}P]dATP from Amersham Corp. Biotin-maleimide (M-1602), BODIPY Labeling Kit (F-6096), TMR-maleimide (T-6027), Lucifer Yellow Iodoacetamide (L-1338), and TNP-ATP

were purchased from Molecular Probes, Inc. Europe (Leiden, The Netherlands). Myricetin, leucocyanidin and tetracycline hydrochloride were purchased from Aldrich; Hesperetin from Sigma; the other chemicals used for inhibitor screening were from Lancaster. Protein broad marker for SDS-PAGE was purchased from Boehringer Mannheim. High protein marker used for native PAGE was from Bio-Rad. T4 Polynucleotide kinase was from New England Biolabs GmbH (Frankfurt, Germany) .

Concentrations of nucleotides were measured spectrophotometrically based on the extinction coefficients of 26400 M^{1} cm^{-1} at 408 nm (TNP-ATP) and 15400 M^{-1} cm^{-1} at 259 nm (other nucleotides).

Oligo(dT) with 4 - 30 nucleotides were synthesized at the Department of Biochemistry of FU Berlin. The concentrations were determined spectrophotometrically using an extinction coefficient of ε_{260} = 8100 M^{-1} cm^{-1} per nucleotide of oligo(dT).

2.2 Methods and Theory

2.2.1 Purification of RepA protein

The RepA protein was prepared using an overproducing strain of *E. coli* (Scherzinger et al., 1991) and purified at 4°C as described (Röleke et al., 1997). The final purified fractions after gel filtration on Pharmacia Sephacryl S200 were analysed by native polyacrylamide gel electrophoresis (without SDS) run at 4 °C with 14 V/cm. The concentration of protein was determined spectrophotometrically using an extinction coefficient of ε_{280} = 25180 M^{1} cm^{-1} (monomer) (Niedenzu, 1997).

2.2.2 Site-directed mutagenesis

Site-directed mutagenesis of the RepA gene was performed by using a PCR technique as described (McPherson et al., 1992). Mutant genes were sequenced to confirm the desired mutation and the absence of spurious mutations that may have been introduced during PCR. K43A mutant was prepared in the same way as described for the wild type and was more than 95% pure as estimated from Coomassie-stained SDS/PAGE gels.

2.2.3 Steady-state kinetics of ATPase assays and determination of inhibition constants

ATP hydrolysis was monitored by measuring the production of inorganic phosphate using acidic ammonium molybdate and malachite green (Lanzetta et al., 1979; Van Veldhoven et al., 1987). Reactions (100µl) were performed at 30°C (unless otherwise indicated). The experiments were started by the addition of RepA at a final concentration indicated (80nM or 160nM hexamer enzyme in the presence or absence of ssDNA) and stopped after 10 min with 800µl 3:1 (v/v) acidic malachite green/ammonium molybdate solution, followed after 1 min by addition of 100µl 34% (w/v) sodium citrate. After incubation for 10min, the absorbance at 660nm was measured using control reactions (without enzyme) as reference. All assays were performed at least in duplicate. The amount of phosphate produced was quantified using a standard curve.

The release rate of inorganic phosphate (Pi) versus ATP concentration (varied at least 5-fold above and below the measured K_M) was used to determine kinetic parameters. They were calculated according to the Hill equation (1) with the nonlinear analysis Origin5.0 Programme.

$$\log \frac{v}{V_{\max} - v} = n \log[S] - \log K_M \qquad (1)$$

where v is the rate of hydrolysis reaction, V_{max} is the maximal rate, [S] is the concentration of ATP, n is the Hill constant.

Inhibition of ATPase activity by ATP analogs was determined using the equation for competitive inhibition:

$$1/v = \frac{1}{V_{\max}}\left[1 + \frac{K_m}{[S]}\left(1 + \frac{[I]}{K_i}\right)\right] \qquad (2)$$

The experimental data for kinetic inhibition by flavone inhibitors were analysed using eq. 3 for noncompetivity (Stryer, 1975):

$$v = \frac{V_{\max}([S]/K_m)}{1 + [I]/K_i + [S]/K_m + S[I]/(K_i K_m)} \qquad (3)$$

where K_m *and* K_i are the substrate binding and inhibition constants, [S] and [I] denote the concentrations of substrate and inhibitor, and v and V_{max} the initial and maximum velocity of the reaction, respectively.

2.2.4 In *vivo* determination of the minimal inhibitory concentration (MIC)

Overnight cultures of *E. coli* SCS1 (Stratagene) or *B. subtilis* SB19 (Nester et al., 1961) were diluted 1000-fold with YT medium. To aliquots of the dilutions, the respective flavones dissolved in DMSO were added. In all assays the overall DMSO concentration was adjusted to 2.5%. Following incubation at 37°C for 18 h, the MIC was determined visually.

2.2.5 DNA unwinding assay (helicase assay)

A forked helicase substrate was used similar to that described by (Crute et al., 1998). To viral M13mp18 DNA, a 5´-^{32}P-labeled 53-mer oligodeoxynucleotide was annealed, resulting in a double-stranded segment of 31 bp and 22 unpaired nucleotides at the 3´ end. Helicase assays were performed in 20 µl buffer C containing 45 fmol helicase substrate, 6 pmol RepA (monomer) and inhibitors as indicated.

In the absence of ATP and helicase substrate, RepA was preincubated with inhibitor in buffer C for 10 min on ice. Then ATP and helicase substrate were added and the mixture was transferred to 30°C. After incubation for 10 min, the reactions were terminated by adjusting the mixture to 25 mM EDTA, 0.25% SDS and proteinase K at 0.1 mg/ml. Following incubation for additional 5 min at 30°C, the displaced 53-mer oligodeoxynucleotides were separated from helicase substrate by electrophoresis on 10% polyacrylamide gels in 89 mM Tris-base, 89 mM boric acid and 1 mM EDTA for 90 min at constant voltage of 120 V. Products were visualized by storage phosphor autoradiography and analyzed with the ImageQuant software version 5.1 (Molecular Dynamics).

2.2.6 Protein-DNA photo-crosslinking

Modified iodo-linked-oligomers $(dT)_{10}$, $(dT)_{20}$ and $(dT)_{30}$ (see Fig. 2.1) were labeled at the 5'-end with [γ-^{32}P]dATP using T4 polynucleotide kinase according to the protocal described by Sambrook et al. (1989). The labeled ssDNAs were mixed with RepA helicase as indicated in buffer A containing 1mM AMP-PNP. After 1/2 hour incubation at 30°C, the mixtures (20µl)

were placed on Parafilm, put on ice, and irradiated for 20 min at a distance of 10cm using a UV lamp with a maximum output of 312nm. After irradiation, the samples were loaded on 10% SDS polyacrylamide denaturing gel and electrophoresis was performed at a constant voltage of 220V. The gels were stained with Coomassie Brilliant Blue, dried, and autoradiographed.

$d(T)_{10}$ 5' - XTTTXTTTXT - 3'

$d(T)_{20}$ 5' - XTTTXTTTXTTTXTTTXTTT - 3'

$d(T)_{30}$ 5' - XTTTXTTTXTTTXTTTXTTTXTTTXTTTXT - 3'

X: 5-Iodouracil T: Thymine

Figure 2.1: Description of the modified iodo-linked-oligomers d(T) used in this study.

2.2.7 labeling ssDNA with BODIPY fluorescence dye

100 µg of ssDNA was phosphorylated at the 5´ terminus with 100 units of T4 polynucleotide kinase in a total volume of 100 µl containing 2.5 mM ATP, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA and 50 mM Tris-HCl buffer pH 7.6. The reaction mixture was incubated for 2 h at 37 °C followed by precipitation of the resulting 5´phosphate terminated ssDNA by ethanol. The 5´-phosphate terminated ssDNA was then dissolved in a volume of 10 µl at a concentration of 10 µg/µl. The 45mer ssDNA used for this study can form partial duplex structure (see Fig. 2.2).

According to the protocol supplied by Molecular Probes Europe, 10 µl of the 5´-phosphate-terminated oligodeoxynucleotide stock solution (10µg/µl) was added $_2O$, 50 µl labeling buffer (0.24 M methylimidazole pH 9.0, 0.32 M EDAC) and 20 µl (10 µg/µl) of the phosphate-reactive cadaverine derivative of BODIPY dye dissolved in DMSO. The reaction mixture was placed on a shaker oscillating at low speed and allowed to react for 4 h at room temperature resulting in reasonably stable phosphoramidate adduct. Finally the BODIPY fluorescence labeled ssDNAs were precipitated with ethanol. The precipitates were dissolved in 50 mM triethylammonium acetate pH 7.0 and applied to reversed phase HPLC using a C18 column with a linear 5%-60% acetonitrile/water gradient over 120 min to separate unlabeled ssDNA, labeled ssDNA and free dye.

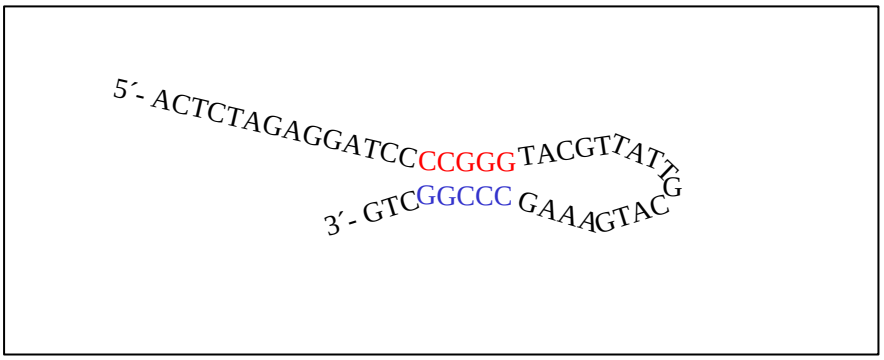

Figure 2.2: The possible secondary structure of 45mer ssDNA used for this study.

2.2.8 labeling RepA helicase with Biotin, TMR and Lucifer Yellow fluorescence dye

In this study, RepA was directly labeled at the unique cystein in the sequence (Cys172) with Biotin-maleimide, TMR-maleimide and Lucifer Yellow Iodoacetamide respectively. To 100 µl solution of RepA (15 mg/ml) containing 20 mM Tris-HCl pH 7.5, 0.1 mM EDTA and 10 mM NaCl was added a 10 µl aliquot of a 10 mM stock solution of TMR-maleimide dissolved in DMSO. The reaction was terminated after 1 h at 25°C by the addition of a tenfold excess of N-acetylcystein with respect to TMR-maleimide. Labeled RepA was passed through a NAP-10 column (Pharmacia) followed by extensive dialysis to remove unreacted dye. The stoichiometry of the labeled product was 1.2 TMR/RepA hexamer as determined by UV-VIS spectrometry. The same method is used for RepA labeled with Biotin-maleimide.

The procedure of labeling Lucifer Yellow Iodoacetamide with RepA is the same as TMR-maleimide except 10 mM stock solution dissolved in H_2O instead of DMSO.

2.2.9 Biotinylation experiment of RepA

The successful biotinylation of a protein is shown by the biotin/streptavidin interaction. Biotinylated helicase was adsorbed on a glass surface. The free binding sites on the glass surface were coated with BSA. In the next step streptavidin (0.5 µM) was bound to the biotinylated helicase. To the remaining binding positions on the streptavidin tetramer, biotinylated actin filaments (200 µM) are bound. The filaments can be detected by fluorescence microscopy.

In the control measurement without helicase, 100 nM Biotin-maleimide and 250 nM acetylcystein were included. This control should make sure that impurities of free biotin in the helicase preparation do not immobilise actin filaments.

2.2.10 Circular dichroism and thermal unfolding measurements

CD spectra were taken at 25°C (unless otherwise indicated) with a Jasco Model 600 recording spectropolarimeter that was thermostated using a constant temperature bath (Lauda). Measurements were performed in the far UV (190-250 nm) range with 0.02cm and 0.1cm pathlength quartz cuvettes to avoid too high absorbance of the sample, and in the near UV (250-320 nm) range 1cm pathlength cuvettes were used. The band width was 1nm and spectra were averaged over at least four scans in order to improve the signal-to-noise ratio. The molar ellipticity [θ] was calculated according to: $[\theta] = 100\ \theta / c\ l$, where θ is the measured ellipticity in degrees, c is the protein concentration in moles per liter, l is the pathlength in centimeters. The amount of secondary structure was calculated with the commercial programme SELCON3 (Sreerama, 1993).

Thermal denaturation was monitored by changes in CD spectra at 222nm as a function of increased temperatures from 25°C to 85°C in a 0.1cm pathlength cuvette. Prior to taking a reading, the solutions were equilibrated for 3min at each newly established temperature. The unfolding curve was analysed with the nonlinear Origin Boltzmann Programme to derive melting temperature T_m and conformational stability ΔG (25°C) (Pace, 1986).

By assuming a two-state reversible equilibrium process between the native (F) and denatured states (U), the fraction of denatured protein f_U at temperature T was determined according to: $f_U = [(\theta_{T,222} - \theta_F) / (\theta_U - \theta_F)]$, where $\theta_{T,222}$ is the measured ellipticity at 222 nm at temperature T, and θ_F and θ_U are the ellipticities at 222nm of totally folded and unfolded states measured at 25°C and 85°C, respectively. The equilibrium constant for denaturation is obtained by: $K = f_u / f_F = f_u / (1- f_u)$, the stardard expression for the temperature dependence of ΔG is: $\Delta G = - RT \ln K = - RT \ln f_u / (1- f_u) = \Delta H - T\ \Delta S$; ΔS is obtained from the slope of a plot of ΔG versus T (K). Since ΔG = 0 at T = T_m, the van't Hoff enthalpy of denaturation at T_m is: $\Delta H_m = T_m \times (\Delta S_m)$. Differences in conformational stability at different experimental conditions are given by: $\Delta(\Delta G) = \Delta T_m \times \Delta S_m$

where ΔT_m is the difference between the T_m values and ΔS_m is the slope of ΔG versus T at T_m in kcal/mol K. In this text, the average slope ΔS_m calculated from all the experimental values is used.

2.2.11 Fluorescence spectrosopy

Fluorescence measurements were performed using a Shimadzu RF-5000 spectrofluorometer. To avoid possible artifacts due to dilution and inner filter effects, the excitation wavelength was set at a spectral bandwidth of 3nm and the emission bandwidth at 15nm (Parker, 1968). Protein concentrations were adjusted as indicated and measured in cuvettes of 1cm pathlength at 25°C (unless otherwise indicated). All titration points were corrected according to equation (4) (Lakowicz, 1983):

$$F_c = (F_i - F_0)(V_i / V_0)10^{0.5d(A_{iex}+A_{iem})} \qquad (4)$$

where F_c is the corrected value of the fluorescence intensity at a given point of titration, F_i is the experimentally measured fluorescence intensity, F_0 is the background, V_i is the actual and V_0 is the initial volume of the sample, d is the total pathlength of the optical cuvette (cm), A_{iex} is the absorption of the sample at the excitation wavelength, and A_{iem} is the absorption of the sample at emission wavelength.

For TNP-nucleotide fluorescence experiments, excitation was set at 410 nm and emission at 534 nm. The solutions were equilibrated for at least half an hour at 25°C before titrations were initiated and stirred with a magnetic micro bar after addition of ligand for about 1min. All titrations were carried out in 1.0 cm pathlength quartz cuvettes. All assays were performed in duplicate.

Fluorescence titration studies on inhibitor binding to RepA were performed at 25°C. The intrinsic binding of inhibitors to RepA was followed by monitoring the quenching of the TNS fluorescence (λ_{ex} = 400 nm; λ_{em} = 440 nm). 2 µM RepA (monomer) solutions and 100 µM TNS in buffer A were equilibrated for 30 min and titrated under stirring in 1 min intervals with aliquots of stock solutions containing inhibitors. The excitation wavelength was set at a spectral bandwidth of 3 nm and the emission bandwidth at 15 nm. 1 ml quartz cuvettes with 1

cm path length were used. All assays were performed in duplicate, titration points were corrected as described above.

2.2.12 Fluorescence correlation spectroscopy (FCS)

For a reversible bimolecular reaction of a labeled small ligand (A*) and a macromolecule (B):

$$A^* \quad + \quad B \quad \Leftrightarrow \quad A^*B$$

the autocorrelation function, $G(\tau)$, can be expressed in terms of a two component model (Meyer-Almes et al., 1998). Since the solution also contains free dye which could not be completely removed by HPLC, it is necessary to apply a three component model according to eq. 5 (Dr. Hecks, EVOTEC GmbH, personal communication). The autocorrelation curves were evaluated with a Marquardt non-linear least squares fitting routine, as implemented in the program FCS Access 2.0 (EVOTEC GmbH). The following three component model corresponding to free BODIPY dye, free BODIPY-ssDNA and the complex between BODIPY-ssDNA and RepA was used:

$$G(\boldsymbol{t}) = \left[1 - P + P\exp\left(\frac{-\boldsymbol{t}}{\boldsymbol{t}_P}\right)\right] * N^{-1}\left[\frac{1-y-z}{\left(1+\frac{\boldsymbol{t}}{\boldsymbol{t}_1}\right)\sqrt{1+\frac{r_0^2}{z_0^2}\frac{\boldsymbol{t}}{\boldsymbol{t}_1}}} + \frac{y}{\left(1+\frac{\boldsymbol{t}}{\boldsymbol{t}_2}\right)\sqrt{1+\frac{r_0^2}{z_0^2}\frac{\boldsymbol{t}}{\boldsymbol{t}_2}}} + \frac{z}{\left(1+\frac{\boldsymbol{t}}{\boldsymbol{t}_3}\right)\sqrt{1+\frac{r_0^2}{z_0^2}\frac{\boldsymbol{t}}{\boldsymbol{t}_3}}}\right] \quad (5)$$

$$with \qquad z = \frac{[A^*B]}{[A^*]+[A^*B]}$$

In this equation, P is the average fraction of dye molecules in the triplet state with relaxation time τ_P; N is the total average number of fluorescent molecules in the observation volume; y and z are the relative concentration fractions of free and bound A*; τ_1, τ_2 and τ_3 define the average time (diffusion time) for detected molecules of free dye, free A* and bound A*, respectively; r_0 and z_0 are the lateral and axial distances between the positions where the Gaussian emission light distribution adopts the maximum value and the point where the light intensity decreases to $1/e^2$ of the maximum value (this defines the observation volume).

Lateral and axial lengths of the observation volume are related through the structure parameter *SP*:

$$z_0 = SP\, r_0 \qquad (6)$$

In order to obtain r_0, the translational diffusion time, τ_{diff}, of a standard (Rhodamine 6G) is measured. The diffusion time is related to r_0 through:

$$\tau_{diff} = \frac{r_0^2}{4D} \qquad (7)$$

where D is the translational diffusion coefficient of the standard. From the experimentally determined diffusion coefficient, an apparent hydrodynamic radius, R_h, can be calculated according to the Stokes-Einstein equation:

$$R_h = \frac{kT}{6\pi\eta D} \qquad (8)$$

were k is the Boltzmann constant, T the absolute temperature and η the viscosity of the solution; at low concentration of solute and buffer, the solvent viscosity can be used.

FCS measurements were performed with a ConfoCor I fluorescence correlation spectrometer (Carl Zeiss Jena GmbH, Jena, Germany; EVOTEC Biosystems GmbH, Hamburg, Germany). The samples were excited using an Argon Laser at a wavelength of 488 nm for BODIPY adducts. Fluorescence emission was detected between 520 nm and 570 nm. The fluorescence of TMR labeled RepA was excited at 514 nm with the same laser, and emission was detected between 520 and 610 nm.

For the determination of binding constants, solutions of 15 nM BODIPY-ssDNA were incubated in buffer (A) or (B) with an excess of RepA (between 0.02 µM and 50 µM) at room temperature for at least 15 min. 100 µl of the sample solution were filled into a chambered coverglass (Lab-Tec, NUNC GmbH, Wiesbaden, Germany), which was placed directly above the objective (C-Apochromat 40x/1.2 water immersion) through which the laser beam passed. The same objective served to collect the fluorescence emission. Spectra were sampled for 45 s if not otherwise indicated. Each single measurement was repeated ten times and the results were averaged. The autocorrelation functions of the intensity fluctuations were automatically recorded

on 288 channels which were quasi-logarithmically spaced in time. These channels cover the dynamic range between 200 ns and 3438 s. Note that a strict temperature control is usually not required, because small fluctuations induced by moderate temperature changes can be neglected (Magde et al, 1974).

Prior to the experiments the structure parameter, *SP*, (see eq. 6 and 7) was determined with a standard Rhodamine 6G solution. For the diffusion coefficient of Rhodamine 6G, a value of 2.8 10^{-10} m ^{2} s^{-1} was used (Rigler, 1993). The translational diffusion time constants of free BODIPY, free BODIPY-ssDNA and the BODIPY-ssDNA/RepA complex were measured in independent experiments and served as input parameters for further fitting procedures. According to eq. 5, the relative concentration fractions of free, y, and bound BODIPY-ssDNA, *z*, the fraction of fluorescent molecules in the triplet state, *P*, the diffusion time of triplet states, τ_P, and the total number of fluorescent species, *N*, served as variable parameters. The fraction of triplet states of BODIPY labeled ssDNA and its complex with RepA was between 10 - 20 % for all experiments. This low triplet yield did not influence the accuracy of data evaluation according to eq. 5.

2.2.13 Time-resolved fluorescence spectroscopy and data analysis

If the molecule under study is fluorescent, it is possible to measure the *fluorescence polarization* which results from excitation with polarized light. This involves measuring the polarization of the fluorescence which is emitted vertical (I_{VV}) and horizontal (I_{VH}) to the incident beam. In this case the fluorescence anisotropy is then defined as:

$$r(t) = \frac{I_{VV}(t) - I_{VH}(t)}{I_{VV}(t) + 2I_{VH}(t)} \qquad (9)$$

The fluorescence decay (without anisotropy) is obtained by measuring at 54.7 degrees angle from the excitation polariser (This angle is also called the magic angle). The total fluorescence intensity, F(t), is defined as: $F(t) = I_{VV}(t) + I_{VH}(t) = F_0\, e^{-t/\phi}$

If the chromophore can undergo motion, then the absorption anisotropy or fluorescence polarization can be used to detect this motion. In the case of linear dichroism the excitation beam of light is used in a manner that results in the chemical destruction of the chromophore. Consequently, the absorbance in the direction of the excitation beam will be reduced. Due to

molecular motion, a redistribution of bleached molecules will occur as a function of time. The population of molecules which are capable of absorbing polarized light is increased as non-bleached molecules diffuse and obtain the orientation of the bleached molecules. There is no time limitation on this motion thus a very wide range of time scales can be probed. In the case of fluorescence an initial polarization is induced into the sample by the excitation. This polarization is destroyed due to molecular motion. If the rate of this motion is such that rotation of the chromophore can occur during the lifetime of the excited state (i.e. before the fluorescent photon is emitted) then the direction of the emitted light will be different than what it would be for an chromophore which is not mobile. This rotation results in depolarization of the emitted fluorescence. Since the lifetime of the fluorescent state is on the order of nsec, fluorescence depolarization can be used to measure rotations which occur in the nsec time frame. Motions which are much faster than a nsec lead to complete depolarization while motions which are much slower than a nsec result in perfectly polarized fluorescence emission.

Fluorescence decays were measured employing a tunable laser/microchannel plate based home-built apparatus for time-correlated single-photon counting with picosecond time resolution. The sample was excited at 430nm with the second harmonic output (LBO crystal, Frequency Doubler model #3980, Spectra Physics) of a frequency-doubled diode-pumped Nd:YVO_4 laser-pumped mode-locked Ti:sapphire laser (Millenia Vs and Tsunami, Spectra Physics) with a pulse width of 1.5 ps (FWHM) at a repetition rate of 4.05 MHz. A fraction of the linearly vertically polarized light was directed onto a nanosecond-photodiode (DET210, Thorlabs). The negative amplified photodiode pulse was passed into a constant fraction discriminator (model #TC 453, Tennelec) whose output in turn after passing through a delay line (model #425A, EG&G Ortec) was used as the stop input signal for the biased time-to-amplitude converter (TAC, model #457, EG&G Ortec). The excitation light pulse hits the sample in a light tight box. The fluorescence was collected at right angles. The cuvette holder for the sample (thermostated) and the reference were mounted on a motorized micropositioner, which was aligned in 45° angle with its axis to the excitation and emission path. This geometry enables the positioner to set either the cuvette containing the sample or the scattering solution (for measuring the instrument response profile L(t)) into the optical path at the same defined and reproducible position. The fluorescence emission was detected

after passing through a cut-off color glass filter model # GG 475 nm for the sample containing Lucifer Yellow or a suitable neutral density filter for the reference, a lens, an iris diaphragm and a sheet polarizer (Pol. UV2, LINOS) onto a microchannel plate photomultiplier tube (MCP) (model #R3809U, Hamamatsu). The parallel ($I_{VV}(t)$) and perpendicular ($I_{VH}(t)$) components for the fluorescence anisotropy decay measurements were detected separately by orienting the sheet polarizer vertically and horizontally against the vertical orientation. The MCP operated at –2900V. The dark counts were reduced to less than 10 Hz using a thermoelectric cooled housing (model # TE-177RF, Products for Research, Inc.). The MCP output was amplified with a 1 GHz preamplifier (model # 9360, EG&G Ortec) and passed into a pico-timing discriminator (model #9307, EG&G Ortec) whose output served as the start pulse for the TAC. The ratio of the stop/start pulse rate was normally less than 200:1, appropriate for single photon counting. The TAC output was transferred into an analog–to-digital converter (ADC, model #8715, Canberra). The ADC data output provides the signals for a computer based multichannel analyzer (MCA) (Accuspec/B, Canberra). The channel width of the MCA was 17 ps and data were collected in 1024 channels. The instrumental response function (IRF) was determined at the corresponding wavelengths with LUDOX (Grace), a colloidal silica solution, as scattering material. The IRF of the system was typically about 40 ps (FWHM). The fluorescence decay profiles ($I_{VV}(t)$ and $I_{VH}(t)$ or $I_{VV}(t) +2I_{VH}(t)$) and the time resolved anisotropy as given by eq. 9 were analyzed with a PC using the software package Global Unlimited V2.2 (Laboratory for Fluorescence Dynamics, University of Illinois). The time course of the fluorescence was fitted with a sum of exponentials

$$I(t) = \sum_{i=1} \alpha_i e^{-t/\tau_i} \qquad (10)$$

The anisotropy decay was fitted with the model function

$$r(t) = \sum_{i=1}^{n} \beta_i e^{-t/\rho_i} \qquad (11)$$

For the hexameric DNA-helicase RepA the rotational correlation time (ϕ) due to the tumbling of the whole complex is on the µs time scale (~300 µs) and therefore out of the fluorescence lifetime scale (nanaoseconds) of the label used. Therefore, the anisotropy decays virtually to a

constant end value r∞. The value of r∞ represents a measure of the degree of sterical hindrance by the protein surface.

2.2.14 Photon correlation spectroscopy (PCS)

In a typical light scattering experiment, a laser beam impinges on a solution and the scattered light is recorded by a photomultiplier. The spatial resolution of the experiment is defined by the scattering vector q whose magnitude is given by the Bragg formula:

$$|q| = \frac{4\pi n}{\lambda}\sin\left(\frac{\theta}{2}\right) \qquad (12)$$

here λ denotes the wavelength of the scattered light, n the refractive index of the solution and θ the scattering angle.

In a PCS experiment the fluctuations of the scattered light due to the Brownian motion of the particles are analyzed in terms of the autocorrelation function (ACF), $G^{(1)}(\tau)$, which is proportional to the distribution of relaxation times, τ_{rel}, and scattering amplitudes of the examined components:

$$G^{(1)}(\tau) \propto \int_{R_{min}}^{R_{max}} N(R)M^2(R)P(q)S(q)\exp(-mR^{-1}q^2\tau_{rel})dR \qquad (13)$$

here m is a proportionality constant, $N(R)$ and $M(R)$ denote number and mass of particles with radius R between the integration limits R_{min} and R_{max}, P(q), and S(q) denote form and static structure factors of the particles. If the particles are non-interacting and small compared with the employed wavelength, the translational diffusion coefficient, D, can be determined through Laplace inversion of the ACF. From D the apparent hydrodynamic radius, R_h, of the particles are calculated according to the Stokes-Einstein equation (eq. 8).

PCS experiments were performed using a light scattering facility (Dierks and Partner, Hamburg, Germany). Measurements were performed at a scattering angle of 90°. A 50 mW diode laser (687 nm) served as light source. Intensity ACFs were automatically accumulated for 20 s in the relaxation time range between 1 µs and 10 s. Under these conditions, particles in the size range between 2 nm to 1µm can be detected simultaneously. ACF spectra were Laplace inverted by CONTIN (Provencher, 1982) which is implemented in the facility used.

For light scattering experiments, the samples were diluted with 40 mM Tris/HCl buffer of pH 7.6 (6.5) or 40 mM MES/NaOH buffer of pH 5.8 to a final protein concentration of 0.15 mg/ml. In order to remove dust, all samples were filtered before each measurement through sterile filters (Minisart; Sartorius, Göttingen, Germany) with 200 nm or 800 nm pore size.

2.2.15 Analytical ultracentrifugation

Molecular mass determination: In order to study RepA under different pH and salt conditions, ultracentrifugation experiments were performed using an XL-A analytical ultracentrifuge (Beckman, Palo Alto, CA) equipped with absorbance optics. For determination of the molecular mass, M, of RepA, the sedimentation equilibrium technique was applied which allows to determine M directly according to eqs. 14 and 15.

$$c_{(r)} = c_0 e^{MF} \tag{14}$$

with

$$F = \frac{(1 - \rho\bar{\nu})\omega^2 (r^2 - r_0^{\,2})}{2RT} \tag{15}$$

ρ is the solvent density, $\bar{\nu}$ the partial specific volume of RepA, ω the angular velocity, R the gas constant and T the absolute temperature, $c_{(r)}$ the radial concentration and c_o the corresponding value at the meniscus position.

In addition to the molecular mass obtained, the determined parameters allow the calculation of the volume,V, of the solute:

$$V = \frac{M\bar{\nu}}{N_A} \tag{16}$$

with N_A Avogadro´s number.

Assuming that a protein molecule has spherical shape, the radius, r, of the unhydrated protein is given by:

$$r = \sqrt[3]{\frac{3V}{4\pi}} \qquad (17)$$

In aqueous solution, the radius is larger by about 0.3 nm (water shell).

Complex formation between RepA and $(dA)_{30}$:

Complex formation between different macromolecules can be easily detected from mass determinations when the associates are formed with high affinity ($K_a > 10^8\ M^{-1}$). Because complex formation is often weak, it is necessary to estimate the association constants of an interacting system by fitting the sum of exponential functions, given in eq. 18, to the experimentally obtained radial scanning curves (Behlke et al., 1995).

$$A_r = \varepsilon_R c_R e^{BM_R F} + \varepsilon_L c_L e^{M_L F} + c_R \sum_{j=1}^{n} (\varepsilon_R + j\varepsilon_L) c_L^j K_j e^{(BM_R + jM_L)F} \qquad (18)$$

Where ε_R, ε_L, c_R and c_L are the extinction coefficients and concentrations of the free protein molecule (R = RepA) or the free ligand concentration (L = $(dA)_{30}$) at the radial position r_i, respectively. B means the difference in buoyancy between R and L, j the number of possible binding steps, and K_j the binding constant corresponding to one binding step. To obtain more precice data for the estimated parameters, we have to reduce the number of variables describing the binding reaction (i) by separate determination of molecular masses, (ii) by using a statistical binding model for equal binding sites (Wyman et al., 1999), and (iii) by taking the mass conservation into account. In each experiment, three absorbance profiles were measured at three different wavelengths. This allowed to determine the total concentration in an arbitrary sector by numerical integration. According to a statistical model, K_j is given by:

$$K_j = (1/n^j)\binom{n}{j} K_1^j \qquad (19)$$

where K_1 is the binding constant for the first step. If one substitutes $(1/n^j)$ $\binom{n}{j}$ with G_j and $1/K_1$ by K_d or $c_L K_1 = \chi$, equation 18 can be written as follows (Behlke et al., 1997):

$$A_r = \varepsilon_R c_R e^{BM_R F} + \varepsilon_L K_d c e^{M_L F} + c_R \sum_{j=1}^{n} (\varepsilon_R + j\varepsilon_L) G_j c^j e^{(BM_R + jM_L)F} \quad (20)$$

For the total concentrations c_{Rt} and c_{Lt}, the integration of the model function results in eqs. 21 and 22:

$$c_{Rt}(r_b - r_m) = c_R \left[\int_{r_m}^{r_b} e^{BM_R F} dr + \sum_{j=1}^{n} c^j G(j) \int_{r_m}^{r_0} e^{(BM_R + jM_L)F} dr \right] \quad (21)$$

$$c_{Lt}(r_b - r_m) = \left[K_d c \int_{r_m}^{r_b} e^{M_L F} dr + c_R \sum_{j=1}^{n} j c^j G(j) \int_{r_m}^{r_0} e^{(BM_R + jM_L)F} dr \right] \quad (22)$$

With r_b and r_m the radius position at the bottom of the cell and the meniscus, respectively. The substitution of c_R and K_d by functions of eq. 21 and 22 allows to reduce the number of estimated parameters to only c. Three of the absorbance profiles (eq. 21 or eq. 22 represents such a profile only for one wavelength) were simultaneously fitted (global) by nonlinear regression using the program "Polymole" which was successfully applied earlier in the analysis of other complexes (Behlke et al., 1994; Behlke et al., 1995; Kraft et al., 2000). The optimal fit to the radial distribution curves allowed to estimate the dissociation constant of the complex.

For determination of the molecular mass, M, the sedimentation equilibrium data were analyzed by means of externally loaded six-channel centerpieces of 12 mm optical path length filled with 70 µl of solution in the corresponding compartment. This cell type allows the analysis of three solvent/solution pairs. The protein concentration was adjusted to 0.16 – 0.48 mg/ml throughout. The sedimentation equilibrium was reached after 2 h overspeed at 14,000 rpm, followed by an equilibrium speed at 10,000 rpm at 10 °C for about 30 h. The radial absorbancies of each compartment were recorded at three different wavelengths, 280, 285 and 290 nm at pH 7.6 and, 275, 280 and 285 nm at pH 5.8. The molecular mass, *M*, determinations were done by simultaneously fitting the three radial absorbance distribution curves according to eqs. 14 and 15 applying a partial specific volume, $\bar{\nu}$ = 0.74 g cm^{-3} which was calculated from the known amino acid sequence of RepA and the density increment.

To study complex formation between RepA and $(dA)_{30}$, radial distribution curves were measured by means of externally loaded six-channel centerpieces of 12 mm optical path length filled with 70 µl of solution in the corresponding compartment. The radial absorbancies of each compartment were measured at three different wavelengths, 280, 285 and 290 nm at pH 5.8 and 275, 280 and 285 nm at pH 7.6 after 2 h overspeed at 14,000 rpm, followed by an equilibrium speed at 10,000 rpm at 10 °C for about 30 h. The protein and oligonucleotide concentrations were adjusted to 0.46 µM RepA dimer and 0 $_{30}$ at pH 5.8, respectively. At pH 7.6 protein and oligodeoxynucleotide concentrations of 1.87 µM RepA and 0 – 6.5 µM $(dA)_{30}$ were used, respectively. The dissociation constants of the complex between RepA and $(dA)_{30}$ and the concentrations of free RepA, free $(dA)_{30}$ and their complexes were determined at pH 5.8 and 7.6 by fitting the three radial distribution curves to eqs. 20, 21 and 22. To calculate the stoichiometries for the binding reaction, the concentration of the complex was determined for different molar ratios $(dA)_{30}$/RepA dimeric hexamer and $(dA)_{30}$/RepA at pH 5.8 and pH 7.6, respectively. The concentration of the complex per RepA dimeric hexamer was plotted as a function of the molar ratio $(dA)_{30}$/RepA dimeric hexamer at pH 5.8. At pH 7.6 the concentration of the complex per RepA was plotted as a function of the molar ratio $(dA)_{30}$/RepA.

2.2.16 Electron microscopy and 3D-reconstruction

Electron microscopy. For electron microscopy investigations (cryo-TEM) and the subsequent 3D-reconstruction 10 µM RepA protein (monomer) were prepared in 40 mM MES buffer solution (pH 5.8) in the presence of 10 mM $MgCl_2$, 60 mM NaCl and 1 mM ATPγS. Droplets of the sample (5µl) were applied to hydrophilised (glow discharge in a BALTEC MED 020 (BAL-TEC AG, Liechtenstein) for 60s at 8W) carbon covered microscopical coppergrids (400 mesh) and supernatant fluid was removed with a filterpaper until an ultrathin layer of the sample was obtained. Again, a droplet of contrasting material (2% Phosphotungstic acid) has been added and blotted. A heavy metal stained layer of the protein assemblies was thus prepared and was subsequently plunged into liquid ethane for vitrification in order to preserve the protein structure in its fully hydrated state. The Gatan-626 specimen holder/cryo-transfer system was used in a Philips Tecnai F20 FEG TEM (FEI Company, Oregon) keeping the sample constantly at a temperature of -179°C. Imaging was performed under low-dose

conditions using a primary magnification of 61702 x at an accelerating voltage of 160kV. The defocus value was chosen to correspond to a first zero of the CTF at ~ 12 Å This specimen preparation procedure was essentially the same described earlier for the 3D-structure determination of influenza HA (Böttcher et al., 1999). An embedding matrix with a somewhat higher contrast compared to the conventional vitreous-ice preparation is obtained allowing the localisation of single molecules despite the relatively high acceleration voltage of the microscope and the chosen "close-to-focus" imaging conditions.

Reconstruction of the 3D structure. Laseroptical diffraction measurements of electron micrographs were performed in order to select "good" micrographs in terms of optically correct phase contrast transfer (absence of aberration, drift, astigmatism etc.). The selected micrographs were digitised using the Heidelberg "Primescan" drum scanner (Heidelberger Druckmaschinen AG, Heidelberg, Germany) at a pixel resolution of 0.68 Å in the digitised images. All image processing was performed using the Imagic-5 software package (van Heel et al., 1996) running on a personal computer cluster under the "Linux" operating system. 6134 molecules were interactively selected and extracted from the digitised micrographs as 400x400 pixel fields. For computational efficiency these single images were interpolated to 1.36 Å / pixel size for all subsequent steps. Using the "angular reconstitution" technique (van Heel, 1987) Euler angles were a posteriori assigned under D6 symmetry restraint and a first rough three-dimensional reconstruction was calculated. Refinements were performed in a iterative process by using forward projections of the preliminary reconstruction as “anchorset” for consecutive steps. Fourier shell correlation (FSC) (van Heel et al., 1986) of two different 3D reconstructions, each of which included half of the final class averages, was done to assess the resolution. The 3ó threshold criterion curve was then corrected by a factor 12 to account for the assumed D6 point-group symmetry (Orlova et al., 1997). For both 3D structures the resolution obtained in the final 3D reconstruction was determined to be ~16.5 Å The three-dimensional alignment of X-ray data of RepA (1G8Y.pdb) and TEM data has been performed in the context of AMIRA 2.3 software.

2.2.17 Crystallization, data collection, structure determination and model refinement

Crystallization. Crystals of RepA and cofactors were grown by vapour diffusion employing the hanging-drop technique at 18 °C. Normally 3 µl of protein solution (17~25 mg/ml) was

mixed with 3 µl resorvoir solution (unless otherwise indicated). For cryo data collection, single crystals were flash-cooled (100K) with liquid N_2 in a rayon loop using cryoprotectant.

Data collection. X-ray diffraction data were collected at synchrotron BESSYII (Berlin, Germany), RSF (Grenoble, France) and DESY (Hamburg, Germany) using MAR 345 image-plates. The data were integrated and processed with DENZO and SCALEPACK (Otwinowski & Minor, 1997).

Structure determination using Molecular Replacement (MR) techniques. All the structures were determined by molecular replacement on the basis of the 2.4 Å resolution model of RepA (Niedenzu *et al*., 2001) using the AMORE programm (Navaza, 1994). Molecular replacement encompasses techniques which are used in macromolecular crystallography to determine the orientation and position of a molecule in the unit cell using a previously solved crystal structure as the 'search model'. The search and target molecules must have reasonable sequence identity (i.e. > 25 %) for there to be a good chance of success. Generally there are two steps in molecular replacement and these are known as the rotation and translation functions.

Solving the target structure ($\underline{x}'$) involves determining a rotation matrix, [C], and a translation vector, $\underline{d}$, to apply to the co-ordinates of the search model, $\underline{x}$ (Fig. 2.3). Hence, $\underline{x}'=[C]\underline{x}+\underline{d}$

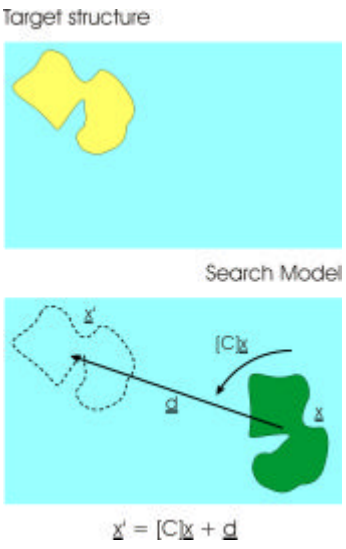

Figure 2.3: Molecular replacement method visually explained. A pictorial representation of the main principle of the molecular replacement method.

This is simplified by separating the search into two stages, namely the rotation and translation searches (divide and conquer).

The rotation function should allow the orientation of the search molecule which produces maximal overlap with the target structure to be determined in the absence of any phases for the unknown structure. To do this, it compares the Patterson self-vectors of the known and unknown structures at different orientations of the search model. Note: The Patterson function can be calculated from the measured structure amplitudes (without phase angles). Using Patterson space means that the translation vector, d, is irrelevant since all intra-molecular vectors are shifted to the origin (Fig. 2.4).

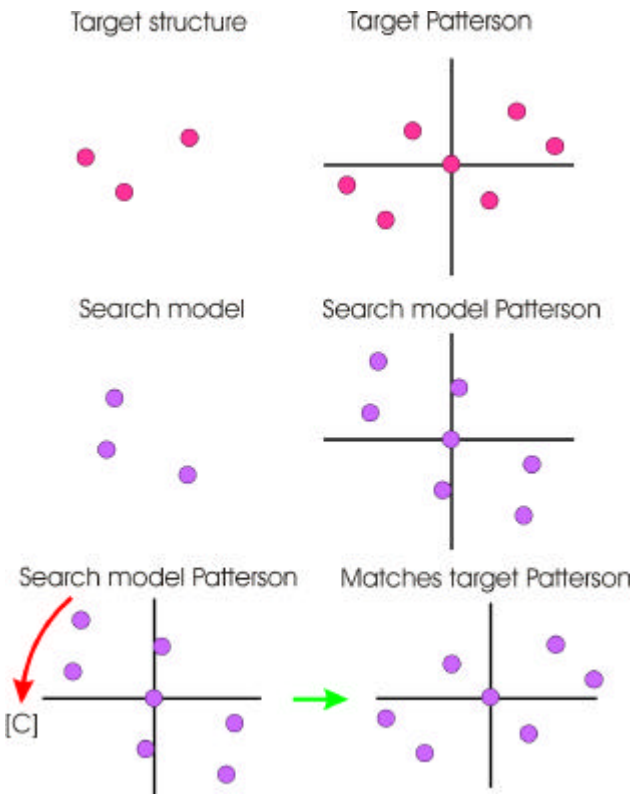

Figure 2.4: Schematic diagram showing the comparations of the Patterson self-vectors of the known and unknown structures at different orientations of the search model.

To quantify the agreement between search and target Pattersons the following function was proposed by Rossmann and Blow (1962).

$$R = \int P_T(\underline{u}) P_S([C]\underline{u}) d\underline{u} \qquad (23)$$

where $P_T(\underline{u})$ is the Patterson function of the target structure and $P_S([C]\underline{u})$ is the Patterson of the search molecule which has been rotated by matrix [C]. The integral is usually calculated for a shell in Patterson space with inner and outer radii limits which allow exclusion of the Patterson origin peak and cross-vectors arising from neighbouring molecules since it is desirable that the rotation function be dominated by self-vectors.

The integral can be calculated numerically in real space but this is slow although it can be speeded up by representing the model Patterson as a 'peak list' (Huber, 1965). Rossmann and Blow (1962) proposed a reciprocal space method for calculating R which is faster. Letting $F_{\underline{h}}$ and $F_{\underline{p}}$ be the structure factors for the target and search structures, respectively, then:

$$R = \int \sum_{\underline{k}} \left|F_{\underline{k}}\right|^2 e^{-2p\, i\underline{k}\underline{u}} \sum_{\underline{F}} \left|F_F\right|^2 e^{-2pi\underline{F}[C]\underline{u}} d\underline{u} \qquad (24)$$

This can be simplified to:

$$R = \sum_{\underline{k}} \sum_{\underline{F}} \left|F_{\underline{k}}\right|^2 \left|F_{\underline{F}}\right|^2 \int e^{-2pi(\underline{k}+\underline{F}[C])\underline{u}} d\underline{u} \qquad (25)$$

The integral over volume can be simplified for spherical molecules to:

$$R = \sum_{\underline{k}} \sum_{\underline{F}} \left|F_{\underline{k}}\right|^2 \left|F_{\underline{F}}\right|^2 G_{\underline{k},\underline{F}} \qquad (26)$$

where $G_{\underline{h},\underline{p}}$ is known as an interference function. $G_{\underline{h},\underline{p}}$ is only large if $\underline{h}$ and $-\underline{p}[C]$ are similar, i.e. only those points that are close together in the two reciprocal lattices being compared are significant in the double summation. The interference function is irrelevent if $|F_{\underline{p}}|^2$ for an isolated molecule is sampled on a fine grid allowing $|F_{p[C]}|^2$ to be obtained by linear interpolation at each orientation (Lattman and Love, 1969). This has been developed by Crowther (1972) who suggested that the Pattersons can be approximated by spherical harmonics and the rotation function can be calculated from these by FFT. This method has

been developed further by Navaza (1994) in the program AMORE. Another popular program is MOLREP (Vagin and Teplyakov, 1997). These two programs, which are in CCP4, allow structures to be solved automatically even when more than one molecule is present in the asymmetric unit. The likelihood of having more than one molecule per asymmetric unit should always be ascertained beforehand by calculating the solvent content of the crystal (e.g. by the method of Matthews, 1968).

Having determined the angles α, β and γ from the rotation search, the rotation matrix [C] can be calculated and applied to the co-ordinates of the search molecule. The shift vector ($\underline{d}$) which is required to position the search molecule correctly relative to the symmetry elements of the target crystal can be determined by one of a number of translation searches. For triclinic crystals the translation search is irrelevant and similarly for other polar space groups the problem reduces to a lower dimensional search, e.g. in monoclinic cells only a 2-D search on x and z is needed because the position along the y-axis is arbitrary. The translation function can be used to resolve space group ambiguities (e.g. $P4_12_12$ and $P4_32_12$) - the correct one will gives the highest peak.

The translation search measures the overlap of the target Patterson cross-vectors with those calculated for the oriented search molecule as it ranges through the target cell. The translation function is defined as:

$$T(t) = \int P_c(\underline{u},t) P_o(\underline{u}) d\underline{u} \qquad (27)$$

where $P_c(\underline{u},\underline{t})$ is the search molecule cross-Patterson and $P_o(\underline{u})$ is the observed Patterson function of the target structure. Crowther and Blow (1967) expanded this in reciprocal space giving:

$$T_2(t) = \sum_{\underline{k}} I_o(\underline{h})[\sum_i^n \sum_j^n F_m(\underline{h}A_i) F_m^*(\underline{h}A_j) \exp(-2\pi i \underline{h}t_{ij})] \qquad (28)$$

where the summation over i and j ($i \neq j$) is the calculation of the cross-vectors between each pair of search molecules (m) related by the n symmetry operators A_i and A_j ($\underline{t}_{ij}$ is the intermolecular vector). To ensure that the T_2 function is dominated by cross-vectors, the self-

vectors calculated for the search molecules can be subtracted from the target Patterson function by calculation of I_{cross}:

$$I_{cross} = I_o(h) - k\sum_i^n \left|F_m(\underline{h}A_i)\right|^2 \qquad (29)$$

where k is a scale factor ($\Sigma I_{obs}/\Sigma I_{calc}$). In the T_2 function modified by Ian Tickle (TFFC in CCP4), the self-vectors can also be substracted from the search model Patterson to improve the signal-to-noise ratio. Tickle showed that T_2 can be expressed as a Fourier summation and can then be calculated by use of an FFT algorithm (Driessen *et al.*, 1991). Normalised structure factors (E values) or use of a negative temperature factor can improve the signal-to-noise ratio of the translation function. Many molecular replacement programs allow a packing function to be calculated. In MOLREP this automatically down weights translation function peaks which correspond to positions where the molecules would overlap.

The R-factor search used to be a very popular way of solving the translation problem partly because of its conceptual simplicity. It involves the calculation of an R-factor as the search molecule and its symmetry mates are moved through the unit cell of the target crystal. The correct position should give the lowest R-factor. Other parameters, such as the correlation coefficient, can be used to measure the agreement between the F_o's and F_c's as the search model is moved around.

$$R = \frac{\sum_{\underline{k}} \left\| F_{\underline{k}}(obs)\right| - k\left|F_{\underline{k}}(calc)\right\|}{\sum_{\underline{k}} \left|F(obs)\right|} \qquad (30)$$

An approximate scale factor (k or $\Sigma F_o/\Sigma F_c$) can be determined prior to the R-factor search from the vector sum of the starting F_c's. Note that the correlation coefficient is independent of k.

The contribution that each symmetry related search molecule makes to the F_c's can be obtained by calculating a hemisphere of structure factors for a single correctly oriented search model in the target cell assuming *P*1 symmetry. The so-called partial F_c's for the symmetry

related search molecules are obtained from the different asymmetric units of the diffraction pattern. Moving the molecules around can be represented by phase shifting the partial F_c's:

$$F_c = F_{c1} \exp(2\pi i \underline{h}\underline{t}_1) + F_{c2} \exp(2\pi i \underline{h}\underline{t}_2) + \ldots \quad (31)$$

where $\underline{t}_1$ and $\underline{t}_2$ are the shift vectors applied to each symmetry mate of the search model. An R-factor and/or correlation coefficient is calculated for each position on a grid scanning all possible positions in the asymmetric unit. It is sensible to exclude weak reflections from the search due to errors and the small contribution that they make to the R-factor. The main disadvantage of the R-factor search is that it is slower than the T_2 search.

Finally, it is necessary to check that a peak in the translation function does not cause neighbouring molecules to clash. If the solution is still ambiguous it is necessary to rigid-body refine each possible solution and the correct one should produce an interpretable electron density map.

Model refinement. Prior to refinement 5% of the reflections were selected in thin shells evenly over the entire resolution range and set aside. These reflections were used for cross validation of the refinement process by using the R-factor R_{free} (Brünger, 1992). Reflections that contribute to R_{free} were not used throughout the entire refinement process, however they were included for map calculations. The model was subjected to rigid body refinement then improved by simulated annealing followed by energy minimization and B value refinement using CNS (Brünger et al., 1998). For manual rebuilding the computer graphics package O (Jones *et al.*, 1991) was used. Water molecules were added in the Fo-Fc difference electron density map at peaks which are 3 r.m.s. deviation above mean density and within hydrogen bond distances to protein atoms or other water molecules.

2.3 Software

Protein secondary structure estimation program:
SELCON3 (Sreerama, 1993)

Data processing and reduction:

HKL program package (Otwinowski & Minor, 1997)

Various data modification programs:

CCP4 program suite (Collaborative Computational Project, Number 4. 1994)

Refinement:

CNS (Brunger *et al*., 1998)

Model building:

O (Jones *et al.*, 1991)

Graphics:

Molscript 2.1.2 (Kraulis, 1991)

Bobscript 1.4 (Esnouf, 1997)

Raster3D (Merritt & Murphy, 1994)