

Executive summary:

DNA mismatch repair (MMR) plays a crucial role in the maintenance of genome stability. Loss of mismatch repair gives rise to a mutator phenotype. In hereditary non-polyposis colon cancer (HNPCC) families, germ line mutations in MMR alleles predispose to cancer of the colon, endometrium, ovary and other organs. Mismatch repair is a highly complex, but also highly efficient mechanism of DNA metabolism. It improves the fidelity of DNA replication by up to three orders of magnitude, which requires that its constitutive factors cooperate in a precisely defined manner. Despite years of research, the basic mechanistic insights into mismatch repair were limited. In this project, the mismatch2model consortium used a systems biology approach to learn about the rules that govern the steps of MMR.

To adopt and exploit a systems biology approach towards the elucidation of a complex biological process relevant to human health, mismatch2model set itself the following tasks: to gather quantitative and structural datasets describing the principal steps of DNA mismatch repair; to integrate and analyze structural and functional data obtained from single molecule and bulk studies; develop mathematical models for steps in the MMR pathway; perform analyses at different levels of complexity and combine the well-understood MMR E.coli model system with the cancer-protective human system.

The mismatch2model consortium started off by standardizing the creation of reagents and assay protocols, allowing accurate determination of parameters for quantitative modeling. New elegant biochemical assays and exciting sophisticated single molecule techniques were successfully established. mismatch2model then structurally, mechanistically and kinetically described the initial MMR reaction step of mismatch binding. Subsequently, protein complex formation and activation of DNA modification activities was elucidated mechanistically and kinetically, both for the E.coli and human MMR systems. Several of these results represent paradigm shifts within the MMR field. Finally, combining the results and insights from abovementioned experiments, mismatch2model was able to construct mathematical models for MMR at different levels of complexity. These models reproduce in vitro data and thus provide an understanding of the near 100% efficiency of the in vivo systems and have predictive value towards observed phenotypes of HNPCC mutations.

The mismatch2model consortium consisted of six different laboratories located in five countries, all contributing a unique scientific expertise. Within the consortium the level of collaboration and exchange of expertise, reagents and personnel has been extraordinary. Members of mismatch2model disseminated novel findings at scientific conferences. The mismatch2model consortium has been unique in its integration of basic research disciplines devoted to unraveling a multi-component, highly dynamic repair system with relevance to human health and disease.

Project Context and Objectives:

DNA mismatch repair plays a crucial role in the maintenance of genomic stability. In *E. coli*, the repair process is initiated by the binding of a MutS homodimer to mismatches or unpaired nucleotides that arose as errors of DNA polymerase and escaped the proofreading activity of the DNA replication machinery or during recombination between homologous but non-identical DNA sequences. Mismatch binding triggers an ATP-dependent conformational change in MutS, which then recruits a MutL homodimer. The formation of this ternary complex subsequently activates a cascade of events that involve the recruitment of the strand-discrimination endonuclease MutH, DNA helicase II (UvrD), one of several exonucleases, DNA polymerase III and DNA ligase.

The main features of the DNA mismatch repair system are conserved from *E. coli* to man. In eukaryotes, the MutS function is mediated by MutS homologues (MSH); the heterodimer of MSH2 and MSH6 (MutS α) is mainly involved in the recognition of base-base mismatches, while MutS β (a heterodimer of MSH2 and MSH3) is primarily involved in the recognition of small insertion and deletion loops. Similarly, the eukaryotic MutL homologues function in the form of heterodimers; MutL α is composed of MLH1 and PMS2, MutL β of MLH1 and PMS1, and MutL γ consists of MLH1 and MLH3. There is no MutH homologue in eukaryotes, but recent work suggests that the endonuclease activity resides in the MutL homologues.

In both bacteria and eukaryotes, the loss of mismatch repair gives rise to a mutator phenotype. In hereditary non-polyposis colon cancer (HNPCC) families or Lynch syndrome, germline mutations in one MSH2, MSH6 or MLH1 allele predispose to cancer of the colon, endometrium, ovary and other organs. HNPCC is the most frequent form of familial cancer. The young age of onset and high disease penetrance indicates that the loss of the wild type copy of the mutated gene occurs with high frequency in epithelial cells. Moreover, about 10% of colon cancers display a phenotype identical to that of HNPCC tumours. This phenotype is not linked to MMR gene mutations; rather, it is associated with the epigenetic silencing of the MMR gene MLH1.

Mismatch repair is a highly-complex, but also a highly-efficient mechanism of DNA metabolism. It improves the fidelity of DNA replication by up to three orders of magnitude, which requires that its constitutive factors cooperate in a precisely defined manner. In particular, the steps of mismatch recognition and strand discrimination have to be strictly controlled, given that MMR has to be directed to the newly-synthesized DNA strand (in the case of replication) or to the invading DNA strand (in the case of recombination). How this is accomplished in the eukaryotic system is currently unclear. The *E. coli* system involves the MutH endonuclease, which is activated by the MutS/MutL complex to cleave the unmethylated strand at GATC sites in the newly-synthesised DNA, but it is not known how. Given that the strand discrimination signal can be as far as a kilobase from the mismatch and still direct the repair process, the molecular mechanism of this key signal transmission process is of substantial interest. Indeed, proposed mechanisms for the manner in which MutS, MutL and MutH structurally and biomechanically couple mismatch recognition to strand discrimination vary broadly, from sliding clamp mechanisms to translocating molecular motors. Clearly, the structural and mechanistic details that govern this process are crucial to identifying and understanding the system's rate-limiting features.

In recent years, the field has been able to acquire considerable knowledge about the eukaryotic MMR system. This work culminated in the reconstitution of the human system from purified recombinant constituent proteins. However, as in the case of *E. coli*, we currently lack the basic mechanistic insights into mismatch repair. In this proposal, we plan to use systems biology to learn about the rules that govern the principal steps of MMR: mismatch recognition, repairosome assembly, strand discrimination and strand degradation.

At the present time, the term "systems biology" is most often applied to describe high throughput proteomic studies that attempt to construct protein interaction maps. While these networks contain extremely valuable information, they do not tell us when these protein complexes form, when they disassemble and how these events are regulated. This information is, however, vital to our understanding of the functioning of a cell. Addition of the dynamics and regulation of protein complex formation to the existing interaction networks requires acquisition of affinity constants for the protein partners in the complexes, rates of association and dissociation, and information about the factors influencing these parameters. These kinds of data are generally lacking in systems biological descriptions of cellular pathways. An illustrative example is the Reactome knowledge base (see <http://www.reactome.org> online), in which key biological pathways are depicted step by step. This database contains detailed information on which protein interacts with which partner and how this leads to the next step in the pathway, but there is no quantitative information describing the dynamics of these processes.

Recent EU-efforts (3D-repertoire, Spine2complexes) are aimed at studying protein complexes at the structural level. In an attempt to reach the next level of understanding of complex biological processes, we now study the individual steps of the pro- and eukaryotic mismatch repair process in fine detail, using a combination of high-resolution structural and kinetic experimental approaches with in vitro repair assays and mathematical modelling. This will involve single-molecule force spectroscopy and imaging (DNA nanomanipulation, single-molecule fluorescence energy transfer (FRET), scanning force microscopy (SFM) imaging), surface plasmon resonance imaging (SPR), structural biology (X-ray elucidation of DNA/protein and protein/protein complexes of increasing complexity, crosslinking studies) and mathematical modelling of first the kinetics and energetics of the latter processes, and, second, of the in vivo characteristics of MMR and its interactions with other DNA repair pathways. Thus, this consortium generates, integrates and organizes mechanistic information on a poorly-understood, yet indispensable process of DNA metabolism in both structural and kinetic detail, from the atomic level to the macromolecular and cellular levels. Such an approach at multiple levels is unique and provides a prototype for other cellular processes.

The combination of the individual strengths of the multidisciplinary groups involved in this program ensured that significant inroads are made into the understanding of this complex process. Additionally the program provides insights that will be of relevance to other biological pathways.

CONCEPT

To adopt and exploit a systems biology approach towards the elucidation of a complex biological process relevant to human health. This required multidisciplinary, acquisition and analysis of quantitative data sets

and development of robust mathematical models. In this project we combined Europe's expertise in DNA mismatch repair with sophisticated multidisciplinary technology and expertise in quantitative modeling in order to describe this DNA repair process at different levels of complexity, in an attempt to gain a better understanding of the role of this process in cancer development.

OBJECTIVES

The multidisciplinary network of experts in structural biology, biochemistry, biophysics, mathematical modeling and cell biology will have the following tasks:

1. Gather quantitative and structural data sets describing the principal steps of DNA mismatch repair
2. Integrate and analyze structural and functional data obtained from single molecule and bulk studies
3. Develop mathematical models for single and/or multiple steps of the MMR pathway
4. Perform analyses at different levels of complexity: from recognition of the mismatch through repairosome assembly to in vivo function
5. Combine the well-understood MMR E.coli model system with the cancer-protective human system

These objectives relate directly to the topics addressed by the call HEALTH-2007-2.1.2-5:

Projects should be multidisciplinary and should focus on collecting, analysing and applying quantitative data to enable system biological approaches addressing basic biological processes at all appropriate levels of system complexity.

To achieve the abovementioned aims, the activities within the project have been organized in Work Packages as follows:

WP1: DEFINITION OF CONDITIONS AND PREPARATION OF REAGENTS

1. To define standard assay conditions and prepare standardized reagents
2. To benchmark functionality of mutant variants and labelled derivatives

WP2: QUANTITATIVE ANALYSIS OF MISMATCH RECOGNITION BY MUTS AND MUTSALFA

1. To provide kinetic and thermodynamic analysis of DNA mismatch binding and release by the MutS/MutSalfa proteins
2. To compare analysis of mismatch recognition at different levels of complexity (single-molecule versus bulk) and purity (purified protein versus cell-extract)

WP3: ASSEMBLY OF THE PRE-INCISION COMPLEX

1. To quantitatively determine the assembly and disassembly kinetics of the successive protein components of MMR on a site of DNA damage
2. To measure the DNA remodelling properties of the protein components of MMR on damaged DNA, and to correlate its kinetics with assembly of the MutS/MutL complex and nucleotide hydrolysis by these ATPase components

WP4: COUPLING TO STRAND DISCRIMINATION SIGNAL AND DOWNSTREAM EFFECTORS

1. To obtain kinetic data to model the assembly of enzymes/proteins downstream of mismatch recognition/verification
2. To obtain kinetic data on enzymatic action on the DNA i.e. strand incision, unwinding and excision

WP5: MISMATCH REPAIR IN VITRO AND IN VIVO

1. To obtain MMR rates in human nuclear extracts and in an in vitro reconstituted system and dependence on protein concentrations and HNPCC mutations
2. To obtain in vivo E.coli MMR rates and their dependence on protein concentrations
3. To study kinetics of human MMR complex formation in vivo

WP6: QUANTITATIVE MODELLING OF MISMATCH REPAIR

1. To establish a database for kinetic data on sub steps in DNA MMR and overall repair rates
2. To construct mathematical models for DNA MMR

Project Results:

WP1: DEFINITION OF CONDITIONS AND PREPARATION OF REAGENTS

The goal of this work package was to define standard assay conditions and to prepare reagents, proteins and mutant and labeled variants under standardized conditions such that experiments performed within the consortium could be compared and would return standardized parameters that could be used for modeling in WP6.

Standard assay conditions were defined at the start of the project and were implemented in all laboratories. This meant that all in vitro biochemistry in ensemble and single molecule analysis as well as reconstitution of portions and complete repair reactions using defined components or cell extracts, were performed in very similar or identical conditions. Only if a certain approach required a deviation from the standard conditions, we would adapt, usually by addition of a stabilizing agent such as detergent, glycerol or BSA which otherwise did not influence the reaction conditions.

Several standard DNA constructs were prepared by different partners within the consortium. Short DNA substrates for SPR analysis, fluorescence polarization, FRET analysis and crystallization were designed, commercially ordered, annealed and purified. A large series of linear DNA substrates containing a single mismatch and GATC sites at different positions was constructed for single molecule DNA nanomanipulation. Another series of circular DNA substrates with a defined mismatch and nicking sites at different positions was constructed for in vitro excision and mismatch repair assays. Based on the same backbone, another series of circles containing a defined mismatch and a varying number of hemimethylated GATC sites at different positions were constructed for daughter strand incision analysis. Protocols are now integrated as routine procedures in the different laboratories. A new protocol for the creation of small circular DNAs was implemented and published (Xiao Y, Jung C, Marx AD, Winkler I, Wyman C, Lebbink JH, Friedhoff P, Cristovao M. Generation of DNA nanocircles containing mismatched bases. *Biotechniques* 51, 259-262 2011).

Standard protein purification procedures were developed for E.coli MMR proteins (MutS, MutL, MutH, UvrD) and human MMR proteins (MutSalfa, MutLalfa, RPA, ExoI, PCNA, RFC). Many variants of these proteins were constructed and purified: single cysteine variants of MutS, MutL, MutH and UvrD that can be site-specifically labeled or cross linked; ATPase and DNA binding mutants of MutL, helicase-deficient variant of UvrD, nuclease deficient variants of MutH, ATPase and DNA binding deficient variants of MutSalfa and MutLalfa.

Many MutS and MutL variants were fluorescently labeled, characterized and used in functional assays in different WPs. The availability of many site-specifically labeled versions of these proteins that are still functional, not only concerning DNA binding but also in downstream functional assays, is a major strength of the consortium.

In conclusion, work in WP1 was successfully completed and contributed strongly to the successful experimental wet lab work carried out in work packages 2-5. Especially the standardization of purification procedures and assay conditions allowed for accurate determination of parameters used for modeling in WP6. Furthermore, the constant exchange of protocols

and materials between different partners was of great value to the sharing of knowledge and expertise within the consortium, and had a large effect on the effectiveness of the project, underscoring the importance of these kinds of consortia.

WP2: QUANTITATIVE ANALYSIS OF MISMATCH RECOGNITION BY MUTS AND MUTSALFA

The specific ability of the MutS homologs to recognize very rare mismatches and other types of damage is the critical first step in the repair process. In this work package we have focused on this initiating step, using a combination of state-of-the art structural, biophysical and biochemical methods.

Kinetic analysis depends on appropriate modelling of the processes involved. Understanding the DNA recognition by the E.coli MutS protein was complicated by the equilibrium between dimers and tetramers. We could overcome this making use of obligate dimers and tetramers that were generated by the Giessen group prior to the start of the project. The work was initiated by a detailed characterization of MutS in these oligomeric states. MALLS analysis of the proteins alone and in complex with DNA was used to define the best dimer mutants, and showing that these are primarily in the dimer state. These analyses then led the way to structural characterization of these states. Small angle X-ray scattering (SAXS) was performed, both with the obligate dimer and the tetramer. The data revealed that the tetramer is primarily in the elongated state.

The existence of the obligate dimer state of MutS allowed for the first time the crystallization of full length MutS. Initially a structure was solved of full-length MutS in complex with mismatched DNA. In this structure the tetramerization domain is positioned adjacent to the ATPase domains. Given the absence of density for the linkers and in conjunction with the SAXS data it was concluded that this domain is most likely flexible in the full-length MutS protein.

More recently a structure was resolved of the full-length protein in the absence of DNA. Excitingly, this revealed that the opening and closing of the clamp domains is mediated by a sharp kink in the long helices. This unusual structural feature was unexpected and has implications for the process of repair.

The obligate dimer and tetramer also allowed a detailed analysis of the binding process. We used SPR by Biacore to analyse these processes. In order to obtain good fits we made use of double strand DNA attached to the chip via a short single strand DNA linker. Validation of the method and of these substrates was performed using fluorescence polarization. There was an excellent correlation between these fundamentally different approaches. Kinetic fitting of the SPR data was analysed with EVAL. Interestingly, when the major peaks were fit they were compared to the equilibrium dynamics and the correlation was excellent, showing the validity of the approach and the quality of the data.

In this system we tested the effects of the type of mismatch to the binding process. Different k_{on} and k_{off} values could be fitted for all the different mismatches. Also an analysis of the sequence context was performed. We then addressed the effects of nucleotides in this process. Interestingly, all variations in off-rates for the different mismatches

all disappeared in the presence of ATP, indicating that the clamp forming step is the most critical one.

We analyzed a series of mutants, initially focusing around the magnesium binding site in MutS. In this analysis we realized that the DNA mismatch repair protein MutS acts as a molecular switch. In a manner that resembles the G-proteins it toggles between ADP and ATP states. MutS mutants with decreased affinity for the metal ion are impaired in fast switching and in vivo mismatch repair (Lebbink JH, Fish A, Reumer A, Natrajan G, Winterwerp HH, Sixma TK. Magnesium coordination controls the molecular switch function of DNA mismatch repair protein MutS. *J. Biol. Chem.* 285, 13131-13141, 2010).

We also addressed the role of nucleotides by mutating separately the two MutS monomers, in the mammalian protein MutS α . Testing for DNA recognition and sliding clamp formation we could uncouple the roles of MSH2 and MSH6 ATPase domains in this transition.

Further analysis of the effects of nucleotide binding on DNA recognition was performed using native mass spectrometry. We were able to resolve simultaneous DNA mismatch binding and asymmetric nucleotide binding on *E. coli* MutS. We developed a likelihood based algorithm capable to deconvolute the observed spectra into individual peaks. The obtained mass resolution resolved simultaneous binding of ADP and AMP.PNP to MutS in the absence of DNA. We could show that mismatched DNA regulates the asymmetry in the ATPase sites, since we observe a stable DNA-bound state containing a single AMP.PNP cofactor. This was the first direct evidence for such a postulated mismatch repair intermediate (Monti MC, Cohen SX, Fish A, Winterwerp HH, Barendregt A, Friedhoff P, Perrakis A, Heck AJ, Sixma TK, van den Heuvel RH, Lebbink JH. Native mass spectrometry provides direct evidence for DNA mismatch-induced regulation of asymmetric nucleotide binding in mismatch repair protein MutS. *Nucleic Acids Res.* 39, 8052-8064, 2011).

In all crystal structures of G:T mismatch-bound MutS, phenylalanine is stacked against thymine. To explore whether these structures reflect directional mismatch recognition by MutS, we monitored the orientation of *Escherichia coli* MutS binding to mismatches by FRET and anisotropy with steady state, pre-steady state and single-molecule multiparameter fluorescence measurements in a solution. The results confirm that specifically bound MutS bends DNA at the mismatch. We found additional MutS-mismatch complexes with distinct conformations that may have functional relevance in MMR. The analysis of individual binding events reveal significant bias in MutS orientation on asymmetric mismatches (G:T versus T:G, A:C versus C:A), but not on symmetric mismatches (G:G) (Cristovyo M, Sisamakias E, Hingorani MM, Marx AD, Jung CP, Rothwell PJ, Seidel CA, Friedhoff P. Single-molecule multiparameter fluorescence spectroscopy reveals directional MutS binding to mismatched bases in DNA. *Nucleic Acids. Res.* 40, 5448-5464, 2012).

During the final year of the project major investments in technical development culminated in the establishment of a new and exciting combination of techniques that allows simultaneous visualization and functional analysis of MMR complexes at the single molecule level. The consortium already successfully analyzed mismatch binding in this setup and is very excited about the array of future analyses of the MMR mechanism that have now been made possible.

In addition we studied MutS sliding clamp formation and dissociation from DNA by placing fluorescent labels on protein and DNA and monitoring FRET between these labels using stopped-flow fluorimetry as a readout for these important steps in the mismatch repair reaction.

Several of the abovementioned experiments delivered parameters that were directly used for constructing a quantitative model for MMR: MutS binding to mismatch ($1/200 \text{ sec}^{-1}\text{M}^{-1}$) from SMN+V, MutS sliding clamp formation ($1/2.3 \text{ sec}^{-1}$) and MutS sliding clamp dissociation from DNA ($1/600 \text{ sec}^{-1}\text{M}^{-1}$) from ensemble fluorescence studies.

In conclusion, this work package has been very successful in meeting its original objectives of providing structural and mechanistic insights into as well as of kinetic description of the first step of the repair reaction, namely recognition and binding of a DNA mismatch, and its subsequent signalling response.

WP3: ASSEMBLY OF THE PRE-INCISION COMPLEX

The overall goal of Workpackage 3 has been to study the assembly of the pre-incision MutSL complex on mismatched DNA in terms of its order-of-assembly, kinetic, and structural aspects. Upon mismatch binding and recognition by MutS, MutL assembles onto MutS and will then recruit MutH to carry out incision of the neosynthesized strand. Thus the objectives of WP3 were to 1) quantitatively determine the assembly and disassembly kinetics of the successive protein components of MMR on mismatched DNA 2) measure the DNA remodelling properties of the MutS and MutL components of MMR, and 3) to study the role of ATPases in MutS and MutL in the processes characterized in the first two objectives.

Biacore measurements of MutL recruitment to MutS provided kinetic insight into the association and dissociation kinetics of MutL and mismatch-bound MutS. Binding was seen to take place with an on-time on the order of seconds, and a very slow dissociation process on the order of minutes for DNA with blocked ends. These results indicate that mismatch-validated MutS recruits MutL and that the two protein components are retained on DNA for long periods on the order of tens of minutes (off-rate $\sim 0.003 \text{ s}^{-1}$). Finally, analysis of affinity constants also indicated an increase in retention on longer DNA relative to shorter DNA molecules

Biacore measurements were confirmed using FRET assays when single-molecule evanescent-wave detection (EWD) was found to be too prone to surface artefacts. The FRET assays indicate that the rate of assembly of MutL onto mismatch-validated MutS is on the order of 0.1 s^{-1} (ie requires approximately ten seconds), and that the MutSL complex thus formed on DNA is retained for timescales on the order of several minutes (off rate $\sim 0.004 \text{ s}^{-1}$, or ~ 4 minutes), corresponding to relatively long periods of time. These results thus confirm the Biacore measurements for MutL assembly on mismatch-bound MutS.

Thus, the first objective has been accomplished using Biacore measurements, which have been further confirmed by FRET measurements, providing values for the kinetics of assembly/release of MutS and MutSL on mismatched DNA and that are used for modelling purposes. Finally, these results have also been verified to hold in the context of the overall incision reaction thanks to bulk order-of-addition assays measuring the incision rate. They are also consistent with results published recently by other laboratories working in the field.

Single-molecule nanomanipulation has been successfully used to characterize the DNA remodelling activity of the MutS and MutSL complexes on mismatch-bearing DNA. Formation of DNA loops on the order of several hundred base-pairs by MutS and MutSL obeys simple thermal statistics (ie based on temperature and DNA bending stiffness), and is thus a random thermal process rather than an active motor-like process. At the same time, looping was strictly dependent on ATP hydrolysis by all components, and the obligate dimer variants of MutS also displayed the ability to potentiate this DNA looping. MutS is capable of driving DNA looping on its own at somewhat elevated concentrations (~10-20 nM), but in the presence of MutL DNA looping is observed at sub-nanomolar concentrations of MutS (eg ranging from 0.1-1 nM). Half of loops have a ~5 s lifetime, and the other half a ~50 s lifetime. Finally, force spectroscopy measurements show that loop size decreases with increasing force. These effects are well-described by theoretical models for DNA looping stabilized by DNA-binding proteins.

SFM imaging of MutSL-DNA complexes provided additional insight into DNA remodelling by MutS and MutL. In addition to confirming the size-distribution of DNA loops observed in the single-molecule nanomanipulation assays, the SFM images also provide a means to identify localization of protein components in the loops, as well as the organization of DNA about the protein components. These observations found that MutSL complexes are localized to the base of loops but with no apparent DNA wrapping about the protein. Furthermore, although proteins at the base of loops were often located near a DNA mismatch they were not systematically located on a DNA mismatch, indicating that the MutSL proteins localized to the base of the loops are mobile along the DNA.

Thus the second objective has been achieved using SFM imaging and single-molecule nanomanipulation experiments; these experiments have successfully characterized the DNA remodelling process and observed it to behave as a thermally-activated process rather than an ATP-fueled translocation process. These observations made it possible to rule out one of the oldest hypothesis pertaining to MMR initiation, namely the active formation of a loop "extruded" by a helicase burning ATP (and in a manner analogous to certain restriction-modification systems). Subsequent single-molecule incision experiments (presented in WP4) then indicated that this thermally-driven DNA remodelling process is not essential to the incision reaction (ie is neither required nor rate-limiting) and thus does not warrant inclusion in the modelling process.

Biacore analysis of MutSL complex formation was repeated using ATPase mutants, showing that both wild-type and hydrolysis-deficient MutS recruited MutL in an efficient fashion. Thus, ATP hydrolysis by MutS is not required for MutL recruitment, and so it is in the ATP-bound state that MutS performs recruitment of MutL. As MutS releases from the mismatch when it exchanges ADP for ATP, this suggests that MutS recruits MutL after release from the mismatch and formation of the MutS sliding clamp on DNA. This clearly places MutL recruitment after mismatch release. In a similar fashion, MutL recruitment by wild-type MutS was assayed as a function of the MutL ATPase. Here a MutL variant unable to bind ATP was found to be recruited by MutS-mismatched DNA as efficiently as the wild-type form of MutL. This indicates that ATP binding by MutL is not required for recruitment by MutS. Overall these experiments help clarify the order of assembly and nucleotide states of MutS and MutL as

they assemble into a complex, with ATP-bound MutS released from the mismatch and thus competent to recruit MutL.

ATPase assays for MutS and MutL were carried out as planned, and determined base-line turnovers for individual protein components in the presence of mismatched DNA to be on the order of 6 min⁻¹ (for MutS) and 0.2-0.3 min⁻¹ (for MutL), confirming the higher hydrolysis rate by MutS. For DNA with unblocked ends, these assays found that addition of MutL to MutS left the MutS hydrolysis rate unchanged, as did addition of hydrolysis-deficient MutL. Experiments from competing laboratories showed an increase in overall hydrolysis upon addition of MutL, provided the DNA had blocked ends, and found that this increase was only observed using wild-type MutL but not hydrolysis-deficient forms of MutL. Thus overall these results indicate that wild-type MutL is required for proper activation of MutS or that MutS activates the MutL ATPase. Thus, although MutL does not require ATP binding to be recruited by MutS, it does require ATP binding (and also hydrolysis) to correctly activate MutS.

Thus analysis of ATPase usage and mutants has shown absence of activity in terms of both DNA remodelling and (in downstream experiments) incision for a range of mutants as well as for wild-type components in the presence of non-hydrolysable ATP analog ATP- γ -S, as seen in both single-molecule and bulk incision assays. Thus the Biacore results on ATP usage have turned out here to provide the best information on the role of nucleotide state on MutSL complexes.

To conclude, Workpackage 3 has been successfully completed, providing essential new kinetic and structural insights into the mechanisms and kinetics of assembly of the preincision complex and the role of the preincision complex in DNA remodelling. Numerous results from this Workpackage represent paradigm-shifts in the field, notably the quantitative characterization of DNA remodelling and, combined with results from Workpackage 4, the realization that DNA remodelling is nevertheless not necessary to the incision reaction. Indeed the model for DNA looping as a mechanism for supporting MMR initiation was long one of the most strongly-favored models in the field. This Workpackage has thus allowed for rational simplification of the mathematical model for MMR, and has also provided essential kinetic parameters for the mathematical model.

WP4: COUPLING TO STRAND DISCRIMINATION SIGNAL AND DOWNSTREAM EFFECTORS

WP4 addresses the coupling between mismatch recognition on the one hand and the strand discrimination signal and initiation of strand excision reaction on the other hand. In *E. coli*, the strand discrimination signal is a GATC site that is transiently hemimethylated and that is nicked on the unmethylated daughter strand by endonuclease MutH which becomes activated by the MutS MutL complex. In eukaryotes the *in vivo* signal is unknown but may involve the latent endonuclease of human MutLa. *In vitro* a nick suffices and directionality of the excision step is conferred by PCNA and the RFC clamp loader complex.

We studied this coupling using sophisticated biochemical (Giessen, Rotterdam) and single molecule analysis (Paris, Rotterdam) and iterated with quantitative analysis of the data (WP6), we were able to distinguish between different proposed coupling models, i.e. we could exclude looping of the DNA but demonstrate that tracking along the DNA backbone in a diffusive manner is one key mechanism of MMR. In addition, our analysis

revealed that none of the models existing at the beginning of this project was able to explain the observed data. Our results (including data from other WPs) demonstrate a new and more central role for MutL in MMR.

The objectives of WP 4 were to obtain kinetic data to model the assembly of enzymes/proteins downstream of mismatch recognition/verification, and to obtain kinetic data on enzymatic action on the DNA i.e. strand incision, unwinding and excision. Both of these objectives were met.

An outstanding question in this process is the communication between the mismatch and strand discrimination signal that can be separated in sequence up to thousands of base pairs. At the beginning of this project current models heavily under debate had to be tested e.g. DNA looping (Paris, Rotterdam) movement along the DNA helix (Paris, Giessen) or protein-DNA filaments (Rotterdam). In this task we studied the controversial step of coupling to the strand discrimination signal. Here we studied the various models in a variety of ways, studying DNA looping (Paris, Rotterdam) movement along the DNA helix (Paris, Giessen) or protein-DNA filaments (Rotterdam), while we also followed the kinetics in bulk (Amsterdam, Rotterdam, Giessen) or at the single molecule level (Paris). Coupling to downstream effectors was then studied in a variety of complementary efforts, to provide a full kinetic description that allowed quantitative models in WP6.

An assay suitable for high-throughput data acquisition using fluorescence-energy-transfer (FRET) to monitor in detail the kinetics of association/dissociation of Muth on a mismatch containing DNA in the absence and presence of MutS and/or MutL was developed using newly designed, fluorescently labelled DNA substrates and proteins (constructed and benchmarked in WP1). The existence and nature of a MutSLH-complex could be demonstrated. Using chemical crosslinking we could show that all three proteins are present in one complex. Using catalytically inactive variants we were able to trap complexes in a state prior to DNA cleavage but the kinetic analysis revealed that mismatch-provoked binding of Muth to the GATC measured by the FRET assay is highly dynamic. The dynamic nature of this interaction was confirmed by Biacore experiments. Using combined data from crosslinking and FRET a model for the MutSLH-complex could be proposed (Winkler I, Marx AD, Lariviere D, Heinze RJ, Cristovao M, Reumer A, Curth U, Sixma TK, Friedhoff P. Chemical trapping of the dynamic MutS-MutL complex formed in DNA mismatch repair in *Escherichia coli*. *J Biol Chem* 286, 17326-37, 2011).. This model did not include the CTD of MutL but recent results from crosslinking and FRET suggest that the CTD of MutL is in close proximity to the clamp domain of MutS.

We used single molecule nanomanipulation to determine how varying the distance between mismatch and strand-discrimination, (GATC) site affects the DNA remodeling properties of the system. Initially we were particularly interested in verifying whether or not the distribution of loops sizes observed during DNA remodeling by SLH is correlated to the distance engineered between mismatch and GATC site. The effect of the methylation state of the GATC site was not determined as it turned out not to be relevant for the overall mechanism but only modulating the cleavage reaction by Muth.

The incision reaction worked well on single DNA molecule and allowed monitoring DNA looping and incision at the same time. Extended systematic analysis of DNA looping (e.g. influence of force and concentration of

proteins and ATP) indicate a passive loop capture mechanism rather than an active loop formation. However, careful analysis of the looping and incision reactions revealed no apparent correlation between these two events which effectively ruled out an essential role for DNA looping for the incision reaction and hence favors models involving movement on the DNA. Moreover, active prevention of DNA looping by extending the DNA by force still did not abolish the incision reaction. More detailed analysis involving systematic variation of the distance between mismatch and GATC site on the incision rate are best explained by a fast diffusive mechanism. These results were also confirmed by the ensemble experiments using circular and linear DNA substrates.

After demonstrating that looping is not required for the incision reaction and that diffusive models involving a MutS sliding clamp closely wrapped around the DNA are likely to be the (most) relevant ones, the influence of roadblocks on the incision reaction was expected to give a simple result, i.e. that roadblocks block DNA nicking. Initial plans to use DNA constructs incorporating physical roadblocks (such as tethered antibodies) at various distances between the mismatch and the GATC site were not used. Instead, we constructed a highly sophisticated and in a sense natural roadblock namely a stalled RNA polymerase that allows monitoring the presence of the roadblock in real-time. Initial data demonstrated a strongly impaired incision reaction on DNA that contains a roadblock between the mismatch and the GATC site in agreement with DNA looping not being required for the MMR incision reaction. A more careful analysis in which DNA looping, supercoils etc. were monitored revealing that incision was impaired but not abolished under conditions where a) the roadblock was present and b) DNA looping could be excluded. To resume, these results indicate that 1) DNA looping/remodelling is not required for the incision reaction and 2) the DNA between the mismatch and the GATC site must be free of obstacles for the incision reaction to take place. Taken together this strongly suggested the pre-incision complex travels along the DNA helix as it communicates from the mismatch to the incision site. Although the preincision complex may simultaneously promote DNA looping as it does so, this looping does not seem to be sufficient for the incision reaction to take place. These findings have important implication for the in vivo MMR- process. Current models (in literature) are not sufficient to explain these data. A revised model taken into account a new role for MutL (and its ATPase) was developed (see WP6).

After establishing suitable quantitative ensemble assays, extended kinetic analysis of the incision reactions with a systematic variation of key parameters such as distance between mismatch and GATC-sites and number of GATC-sites revealed that movement of the MutS(LH) complex on the DNA is rapid and not involving active DNA translocation. This is in agreement with the results from task 4.1.4 (see above). A more detailed analysis (distance between mismatch-GATC-site, concentrations, DNA-end effects, order of addition) yielded in a wealth of kinetic data that were used for modeling the MMR pathway (WP6). Using these incision experiments the pathway could be dissected to reveal the key steps in MMR, i.e. mismatch recognition/switching of MutS to clamp diffusive fast on DNA, MutL recruitment and (fast) MutH recruitment/incision.

We next analyzed the mismatch-provoked unwinding by DNA helicase. Kinetics of mismatch-provoked DNA nicking by MutH and its modulation by UvrD using different DNA substrates with one or more GATC sites revealed that UvrD can take over once a nick is introduced into the DNA thereby

preventing further nicking of the DNA by MutH. For a quantitative analysis of the unwinding reaction instead of the initial assay on circular DNA substrates a FRET-based unwinding reaction for UvrD activation was developed. With this assay we showed that ATP- and DNA binding by MutL are required to activate UvrD. In a mismatch-independent reaction, ATP-hydrolysis by MutL is not required. Additional kinetic experiments of the unwinding/excision reaction with circular DNA containing various numbers of GATC-sites indicated a role for multiple incisions in the excision reaction.

Before this study no systematic analysis of the role of ATPase of MutS and MutL on the incision reaction had been carried out or had been carried out under conditions which are not mismatch specific. We could clearly show that both ATP-binding and hydrolysis by MutS and MutL are required for strand incision. ATP-hydrolysis by MutS is required to allow DNA binding whereas ATP-binding is required to form the active signaling (sliding clamp). Moreover the ATP-concentration dependence indicates no role of ATP-hydrolysis for an active translocation mechanism. Finally, the role of ATP-binding/ATP-hydrolysis and DNA binding of MutL were systematically analyzed using mutant MutL variants. Both ATP-hydrolysis and to less extent DNA binding by MutL are requirements for strand incision to be efficient. The results are in agreement with ATP binding and hydrolysis of MutL being absolute requirements for incision activity.

The formation of the preincision complex using purified human MutSa and MutLa was achieved. Distance dependence between mismatch and nicked site revealed a similar dependence as observed for the E. coli MMR reactions and is in agreement with the diffusive model ("sliding hypothesis) of an mismatch activated MutSL-complex. No effective excision was observed at long distances (greater than 1000 bp), indicating a finite lifetime of the complex. Indeed, as initially suspected the mode of coupling between mismatch recognition and strand discrimination is believed conserved and hence this part of the modelling can be transferred (at least in part) from the bacterial to the human system.

Concluding this work package has been extremely successful in that several very elegant and sophisticated assays were established, and an impressive amount of functional data was obtained, resulting in many new insights into the mechanism of mismatch repair and input parameters for quantitative modeling.

WP5: MISMATCH REPAIR IN VITRO AND IN VIVO

The primary goal of WP 5 was to obtain quantitative and semi-quantitative information about the MMR process in E. coli and, primarily, in the human system. The MMR system is highly conserved in evolution, at least as far as the amino acid sequences of the principal components, the MutS and MutL proteins and their orthologs, are concerned. We were interested to learn whether the functions of the individual proteins and the molecular mechanism of the process are also as highly conserved. The answer to this question is not obvious, given that most bacteria and all eukaryotes do not use adenine methylation in strand discrimination and, correspondingly, lack the MutH protein that plays a key role in this process in E. coli, and that the eukaryotic MutS and MutL orthologs exist in the form of heterodimers.

The first challenge was to set up protocols that would enable us to perform the experiments in a reproducible fashion, and - most importantly

- in a way that would make comparison of bacterial and human systems plausible and useful for the model-building project. To this end, we first standardized the preparation of the various DNA substrates. We now have a robust protocol that works in several laboratories and that gives high yields of supercoiled heteroduplex substrates, which can be used with extracts of bacterial and human cells, as well as with reconstituted systems. Given that these substrates were in the past used for in vivo transfection studies, we now have a system that additionally permits the comparison of in vivo and in vitro data.

Using these substrates, we were able to standardize protocols for the preparation of nuclear extracts of human cells, which were either MMR-proficient or -deficient. In a parallel effort, we optimized expression systems (bacterial and baculovirus-based) that provided us with sufficient quantities of purified recombinant human MMR proteins. We could then show that addition of the latter polypeptides to human extracts lacking these factors fully restored the repair capacity of the MMR-deficient extracts. In a second series of experiments, we were able to generate mutants of the key MMR factors, MutSalfa and MutLalfa, carrying amino acid changes that were either designed, or that were identified in patients afflicted with hereditary non-polyposis colon cancer. We could then identify the enzymatic functions (e.g. ATPase) and protein domains that were essential for MMR. In this way, we were able to confirm the phenotype of the patient mutations and their segregation with the disease.

In addition, the experiments provided us with information required for the eventual modelling of the MMR process, namely, the kinetics and the semi-quantitative and quantitative requirements of the repair pathway.

Armed with the above knowledge, we embarked on the reconstitution of the human MMR system from purified individual components. We first generated all the recombinant proteins (MutSalfa, MutLalfa, EXO1, RPA, PCNA, RFC, pol-delta). This was a very difficult task, given that pol-delta and RFC are four- and five-subunit proteins, respectively, and the production of the proteins was fraught with difficulties (mainly lack of solubility). However, we were able to overcome these problems and assemble the MMR pathway from its constituent proteins. Based on findings obtained in nuclear extracts, we were able to show that the rate-limiting step of MMR was mismatch detection and exonucleolytic degradation of the error-containing strand. We therefore concentrated on this step of the process, given that it is not only the rate-limiting step, but also the key mismatch-dependent one.

In an extensive set of assays, we were able to determine the optimal concentrations of the principal components, and the kinetics and efficiencies of the mismatch-dependent strand degradation reaction. We were also able to re-examine the phenotypes of the most interesting MutSalfa and MutLalfa mutants. These results will form the basis of our effort to model the MMR process.

The most challenging task of this Workpackage was the attempt to visualize the assembly of the MMR repairosome in vivo. The concept required, as its first step, the stable expression of GFP-tagged MMR proteins. At the onset of this project, this had not been achieved. It had been postulated that overexpression of MMR proteins is toxic, possibly due to induction of apoptosis. In our past attempts, we succeeded in obtaining only a single stable cell line, HCT15, which

expressed MSH6, albeit only 10% of wild type levels. In the present work, we decided to attempt to express the proteins from an inducible promoter. Numerous attempts to achieve this goal in MMR-deficient cells failed, so we decided to concentrate on two MMR-proficient cells, 293 FlpIn and U2OS FlpIn. These cell lines stably express the tetracycline repressor and carry in their genomic DNA an integrated copy of the Flt recognition site. Transfection with a plasmid containing a cDNA flanked by two Flp sites, together with a vector expressing the Flp recombinase should result in a stable integration of the cDNA insert between the Flp sites in the genome. Using this approach, we were successful in stable integrating and inducibly-expressing both MSH2-eGFP and GFP-MLH1.

We then tested the proteins for functionality. It is not possible to induce mismatches in vivo, but we learned in previous studies that treatment of cells with the methylating agent MNNG activates MMR and brings about recruitment of MMR proteins to chromatin. Using this simple test, we were able to show that both proteins are indeed recruited to chromatin upon MNNG treatment. We then studied the kinetics of this event and were also able to develop a second test, in which the genomic DNA was damaged by high-energy laser. This treatment has been shown in other studies to induce a variety of lesions in DNA, a subset of which is bound by MMR proteins. In these experiments, we were able to show that MMR proteins are recruited to this type of DNA damage extremely rapidly - already at the shortest time point that we were able to study, the proteins were demarcating the laser lines. It is not known which type of damage the MMR proteins are recognizing in the laser lines, but recent evidence from our laboratory suggests that MMR proteins may be recruited to DNA damage by ubiquitylation-mediated DNA damage signaling. We have identified FANL1, a protein that physically associates with MLH1, but also with members of the Fanconi anemia interstrand cross-link repair machinery, particularly FANCD2, which is activated by ubiquitylation. This suggests a novel, unanticipated role for MMR proteins and opens a new avenue to study in our future efforts.

Last but not least, we also tested MutS variants for complementation of a MutS-deficient *E. coli* strain. This work combined two aspects of MMR research: (i) investigation of structural characteristics of the proteins such as criteria for mismatch recognition, interaction with other partners, and (ii) exploiting the structural conservation of MMR proteins to study the phenotype of HNPCC mutations. Also this task of the Workpackage was successful.

In summary, Workpackage 5 was highly complex, but overall extremely informative. It yielded numerous novel insights into our understanding of this complex, multifaceted pathway of DNA metabolism that will no doubt be invaluable in future investigations, not to mention in the investigation of relationship between genotype and phenotype in HNPCC families.

WP6: QUANTITATIVE MODELING OF MISMATCH REPAIR

In this work package all the data that have been gathered in the previous WPs is brought together into mechanistic and quantitative models for different aspects of the mismatch repair reaction. We considered mismatch recognition, daughter strand incision, and the complete prokaryotic and eukaryotic repair pathways. The objectives (to establish a database for kinetic data on sub-steps in DNA MMR and overall repair rates, and to

construct mathematical models for DNA MMR) have both been achieved. Below we will discuss this in detail.

We originally envisaged setting up a database with parameters for quantitative modeling. During the kick-off meeting, the necessity of a database was discussed in detail, and it was decided that extensive databases would not be required. Modeling was carried out with a limited set of parameters that was organized in a tabular format, as reported in core report sections 6.1. However, a repository for protocols, available reagents, parameters, input constraints, models, and brainstorm and discussion material has been set up in the form of a WIKI. This format allowed all partners access and possibility to edit and update.

We set out to integrate the quantitative data on mismatch recognition into a description of the process of mismatch recognition in a mathematical model. To obtain input parameters, use of the dimer mutant of MutS allowed precise kinetic modeling of mismatch binding, which is not possible for wild type MutS. This allowed detailed analysis of the binding of MutS to different mismatches. Interestingly, while significant differences in mismatch affinity were detected, the significant parameter, which is the off-rate in the presence of ATP, did not vary. Additional SPR analysis, ensemble and single molecule FRET analysis, single molecule nanomanipulation combined with fluorescence, as well as native mass spectrometry analysis, resulted in a detailed model for DNA mismatch binding and its regulation by nucleotides.

We also considered the search for a DNA mismatch as a localization problem and quantitatively modeled cost and gain using various localization strategies. This provided insight into the role of multiple target sites on the chance of reaching a specific site, and the influence of the noisy and crowded cellular environment on the chances of localizing a target site. Especially when considering the ability of MutS to non-specifically interact with perfectly paired DNA, the rate of mismatch localization in such a crowded environment will be low (Bernhardsson S, Mitarai N, Sneppen K. Protein localization with flexible DNA or RNA. PLoS One 7, 2012).

Next we addressed the coupling mechanism between mismatch recognition and initiation of daughter strand incision. This can be considered as the crucial event in correct DNA mismatch repair, since the signal that distinguishes the nascent strand from the template strand is available for a limited time. The progress in this task has been an iterative process between different collaborators in the mm2m consortium. The model we used at the beginning of the project developed into a more defined description as more and more data became available along the course of the project. A crucial observation was made, using single molecule nanomanipulation that allowed analysis of DNA looping and DNA incision by MutSLH on mismatched DNA. These experiments identified the DNA remodelling process as a thermally-activated process rather than an ATP-fueled translocation process. These observations made it possible to rule out one of the oldest hypothesis pertaining to MMR initiation, namely the active formation of a loop "extruded" by a helicase burning ATP. Subsequent single-molecule incision experiments then indicated that this thermally-driven DNA remodelling process is not essential to the incision reaction (ie is neither required nor rate-limiting) and thus does not warrant inclusion in the modelling process.

We could therefore construct a model for strand incision based on diffusive communication between mismatch and strand discrimination sites. We allow MutS to bind and release, and include an 'activation' time which we relate to the ATP-induced conformation change to the sliding clamp state that can diffuse along the DNA. We allow MutL to bind and release from this MutS sliding clamp. We also include an 'activation' time for MutL, which reflects the ATP-induced closure of this protein around the DNA. Only this activated MutL state can recruit MutH and we assign a chance that this complex locates and nicks a GATC site.

We then used knowledge about the input parameters obtained in Wp2-4; MutS binding, activation and release from WP2; MutL binding and release from WP3, MutL activation from WP4; diffusion constant from literature, and performed stochastic simulations. We validated this model and optimized parameters against our functional datasets obtained in WP4 on daughter strand incision. Here we had systematically varied reaction conditions and geometry of the strand incision setup. We varied the distance between the mismatch and the strand discrimination sites, we varied the number of strand discrimination sites, we varied DNA topology, we varied the concentration of protein components to be both in the subsaturating regime in which nicking rates are dominated by binding events, as well as in the saturating regime in which nicking rates are dominated by conformational changes and the search for target sites.

Important and unexpected findings are that a relatively low functional MutS loading rate is essential and that the activated complex easily misses a single hemimethylated GATC site upon first passage. Thus the model predicts that nicking may well occur on hemimethylated GATC sites that are separated from the mismatch by several hundred base pairs.

We next considered the complete prokaryotic mismatch repair pathway and integrated all knowledge that we collected during this project. We now have a model that includes the relatively inefficient step of mismatch recognition within a crowded environment in which a large amount of competitive DNA is present and the slow conformational changes in the MutS and MutL conformational switches. Diffusion from mismatch to GATC site is very fast compared to other steps, and because diffusion is possible before full activation of the complete incision complex, the distance between mismatch and GATC site becomes largely irrelevant.

Finally we modeled the human mismatch repair pathway. In earlier WPs we obtained all the key parameters necessary for semi-quantitative comparison of the human and bacterial MMR systems and the findings confirm that the conservation of the two pathways exceeds by far the similarities on amino acid level, but extends to the individual mechanistic and even kinetic aspects. Our results suggest a relatively low 'functional' association rate of MutS_h onto a mismatch, similar to our findings in the bacterial system. Furthermore our model predicts that the DNA excision step is rate limiting and strongly dependent on MutS_h, maybe because complete removal of the DNA between the nick and mismatch by Exonuclease requires multiple loading events of MutS_h. We consider the following implications for in vivo repair: during replication, the first MutS_h to detect the DNA mismatch may be the one that is travelling with the replicating fork, and therefore may be loaded at a higher rate than what we determined under non-replicating conditions. In fact this would be necessary, because of competition from nucleosome assembly occurring about 250 base pairs after the replication fork has passed. After mismatch detection, nucleosome assembly is delayed

such that a time window for repair is generated, and additional MutSalfa molecules that are not travelling with the replication fork but that are required for excision, are able to bind to the DNA (Schöpf B, Bregenhorn S, Quivy JP, Kadyrov FA, Almouzni G, Jiricny J. Interplay between mismatch repair and chromatin assembly. Proc Natl Acad Sci U S A. 109, 1895-90, 2012).

This work package had been highly successful. We have been able to gather kinetic parameters for the different important steps in DNA mismatch repair for the prokaryotic as well as the eukaryotic system. We constructed models at different levels of complexity; mismatch recognition, daughter strand incision as well as the complete repair reactions. The models reproduce in vitro data, thus provide an understanding of the near 100% efficiency of the in vivo systems and have predictive value towards observed phenotypes of HNPCC mutations.

POTENTIAL IMPACT

IMPACT ON HUMAN HEALTH: Mismatch repair is critical for maintenance of genomic integrity and signals to processes such as somatic hypermutation and the DNA damage response. Hence its impact on cellular integrity and human health is large. Even a heterozygous mutation in mismatch repair genes can cause a severe predisposition to cancer and in addition many somatic cancers also suffer from loss of MMR. The mismatch2model consortium now has a better understanding of the pathway, which will contribute to improved information on cancer development and the modes of loss of genomic integrity.

IMPACT ON QUANTITATIVE AND SYSTEMS BIOLOGY: In this project we used a multidisciplinary approach to gather, analyse and apply data describing the formation of crucial cellular protein complexes and how this is regulated. These data allowed us to construct models describing the DNA mismatch repair pathway not only in terms of which protein interacts with which partner, but how fast the complexes form and dissociate and how this is controlled by signals such as specific DNA structures and cofactor binding. This detailed level of modelling goes beyond the construction of protein interaction maps and required pathway modelling, taking into account protein and DNA mechanics. For the MMR pathway this was now possible as quantitative and mechanistic data became available in the course of the project. This integrated modelling of networks, pathways and mechanisms has impact in the fields of quantitative and system biology.

IMPACT ON THE DNA MISMATCH REPAIR FIELD: For the DNA MMR field mismatch2model provided a detailed understanding of the mechanism of DNA mismatch repair. Controversies centering on the role of ATP hydrolysis and the mode of coupling between mismatch and strand discrimination signal have been addressed, settling some of these controversies and moving the MMR field into a new era.

TECHNOLOGICAL IMPACT: Beyond the importance of elucidating the mechanism of action of fundamental biological systems such as the MMR system, single-molecule nanomanipulation and visualization studies in general represent a powerful new way of analysing protein-DNA and protein-protein interactions. Because they provide the researcher with the ability to perform direct, simultaneous observations of both the length- and time-scales of protein-DNA interactions, these techniques are a radical new departure from the standard biochemical techniques used to study such systems. Results obtained in this manner make it possible to probe the nature and number of mechano-chemically coupled rate-limiting steps in an enzyme reaction, identifying the system's kinetic "Achilles' Heel", to the potential benefit of drug developers.

In this context, coupling robust single-molecule nanomanipulation techniques to single-molecule fluorescence detection provided yet another dramatic improvement in our ability to manipulate and analyse biomolecular systems. This enabled us to correlate the recruitment of labeled proteins with their mechanical action on their substrates taking place within the system during the course of its enzymatic cycle.

Finally, an often-overlooked impact of these approaches is that their robust statistical nature is helping to lay the foundation for a new kind of "quantitative biology". Exposing doctoral and postdoctoral students to these methodologies helps train the new generation of molecular

biologists in integrating knowledge from the various branches of science (biology, computer science, mathematics, chemistry, and physics).

CORRELATING SINGLE MOLECULE AND BULK ANALYSIS: This project correlated data obtained from single molecule analysis with those obtained by bulk measurements. We obtained insight into how well multiple events in single molecule setups are described by the ensemble analysis of the same parameter in bulk readout. These insights are of relevance for anyone working in the fast growing field of single molecule biology.

DISSEMINATION ACTIVITIES AND EXPLOITATION OF RESULTS

Scientific dissemination of results occurred via regular reports to the European Union, publication in peer-reviewed scientific journals and presentation of results at scientific meetings. The mismatch2model participants published in high-impact journals such as Nature and Cell and speak at important conferences for the field, such as the relevant Gordon Conferences and DNA repair meetings but also at specialized meetings focusing on structural biology, single molecule biophysics and biochemistry.

Further dissemination and exploitation was realized within the mismatch2model teaching program. Students, young postdocs but also PI's regularly visited other participating laboratories to learn and apply new protocols and techniques. Our multidisciplinary approach towards understanding of a single system ensured that these visits were not only informative but also productive, as new techniques were taught by characterization of samples brought from the home-lab. In fact this mobility of people was fundamental to the success of this project.

In addition, mismatch2model constructed a project website presenting all information necessary for the project partners, the scientific community and the public. Mismatch2model also produced a brochure describing the aims of the collaborative project which was printed as well as posted on the website.

List of Websites:

<http://www.mm2m.eu>