Mapping the Dimer Interface in the C-Terminal Domains of the Yeast MLH1−PMS1 Heterodimer†

Jenny M. Cutalo,‡,§ Thomas A. Darden,‡ Thomas A. Kunkel,‡,|| and Kenneth B. Tomer*,‡

Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, P.O. Box 12233, Research Triangle Park, North Carolina 27709, Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, P.O. Box 12233, Research Triangle Park, North Carolina 27709, and Curriculum in Oral Biology, University of North Carolina, Chapel Hill, North Carolina 27599

Received July 11, 2006; Revised Manuscript Received October 2, 2006

ABSTRACT: Yeast MutLα is a heterodimer of MLH1 and PMS1 that participates in a variety of DNA transactions, including DNA mismatch repair. Formation of the MutLα heterodimer requires that the C-terminal domains of MLH1 and PMS1 interact in a manner that is not yet fully understood. Here we investigate the interactions involved in heterodimerization. Using protein surface modification and mass spectrometry, we identify numerous lysine residues that are exposed to solvent in monomeric MLH1. A corresponding analysis of the MLH1−PMS1 heterodimer reveals that three of these exposed residues, K665, K675, and K704, are no longer solvent accessible in the heterodimer, suggesting that they are within the dimer interface. We refine secondary structure predictions and sequence alignments of C-terminal residues of seven eukaryotic MutL homologues and then develop homology models for the N- and C-terminal domains of MLH1. On the basis of this information, we present a model for interaction of the C-terminal domains of MLH1 and PMS1.

Eukaryotic MutL proteins operate as molecular matchmakers to coordinate multiprotein pathways that repair DNA replication errors [mismatch repair (MMR)],[1] prevent recombination between sequences with imperfect homology, modulate cell cycle arrest and apoptosis in response to DNA damage, and conduct meiotic recombination (recently reviewed in refs *1* and *2*). Because defects in these cellular processes result in genome instability, carcinogenesis, and infertility, it is interesting and important to understand exactly how MutL proteins perform their many tasks. Eukaryotic genomes encode multiple MutL proteins, including MLH1, MLH2, MLH3, PMS1, and PMS2. These are homologous to bacterial MutL, particularly for the 300−400 amino-terminal residues that form a domain that can bind DNA and can bind and hydrolyze ATP to promote the conformational changes that coordinate various biological functions (recently reviewed in refs *1* and *2*).

While bacterial MutL functions as a homodimer, the eukaryotic MutL proteins function as heterodimers that contain MLH1 in partnership with a second MutL protein. One MutL heterodimer with a particularly prominent role in MMR is MutLα (*3*). Yeast MutLα is comprised of the ∼87 kDa MLH1 protein and the 103 kDa PMS1 protein tightly associated [K_d = 86.9 nM (*4*)] with a 1:1 stoichiometry. Initial studies (*5*−*7*) demonstrated that formation of a functional yeast (or human MLH1−PMS2) MutLα heterodimer requires interactions of the C-terminal residues of both proteins (Figure 1). Insights into the residues required for dimer formation also come from the crystal structure of the C-terminal domain of the bacterial MutL homodimer (*8*, *9*), which led to the proposal that helix C in *Escherichia coli* MutL is particularly important for homodimerization. Unfortunately, no structure for the C-terminus of eukaryotic MutL proteins is available, and the C-terminal residues of MutL proteins are substantially more diverse than their N-terminal residues, thereby complicating predictions based on structure-based alignments. Thus, our current understanding of the identity of residues involved in heterodimerization, and the nature of the dimer interface in MutL complexes, is derived from a relatively small number of studies that have mapped the general location of the interaction and implicated a few residues (see Figure 1, ref *10* and references therein, and details in the Results and Discussion). In an attempt to further investigate this important interaction, we use protein surface modification and mass spectrometry, as well as refined sequence alignments and predictions of secondary and tertiary structures, to provide a refined model for the dimer interface.

† This research was supported by the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health.

* To whom correspondence should be addressed.

‡ Laboratory of Structural Biology, National Institute of Environmental Health Sciences.

§ University of North Carolina.

|| Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences.

[1] Abbreviations: MMR, mismatch repair; MS, mass spectrometry; MALDI, matrix-assisted laser desorption ionization; HPLC, high-performance liquid chromatography; ESI, electrospray ionization; EIC, extracted ion chromatogram.

10.1021/bi061392a
Published on Web 12/08/2006

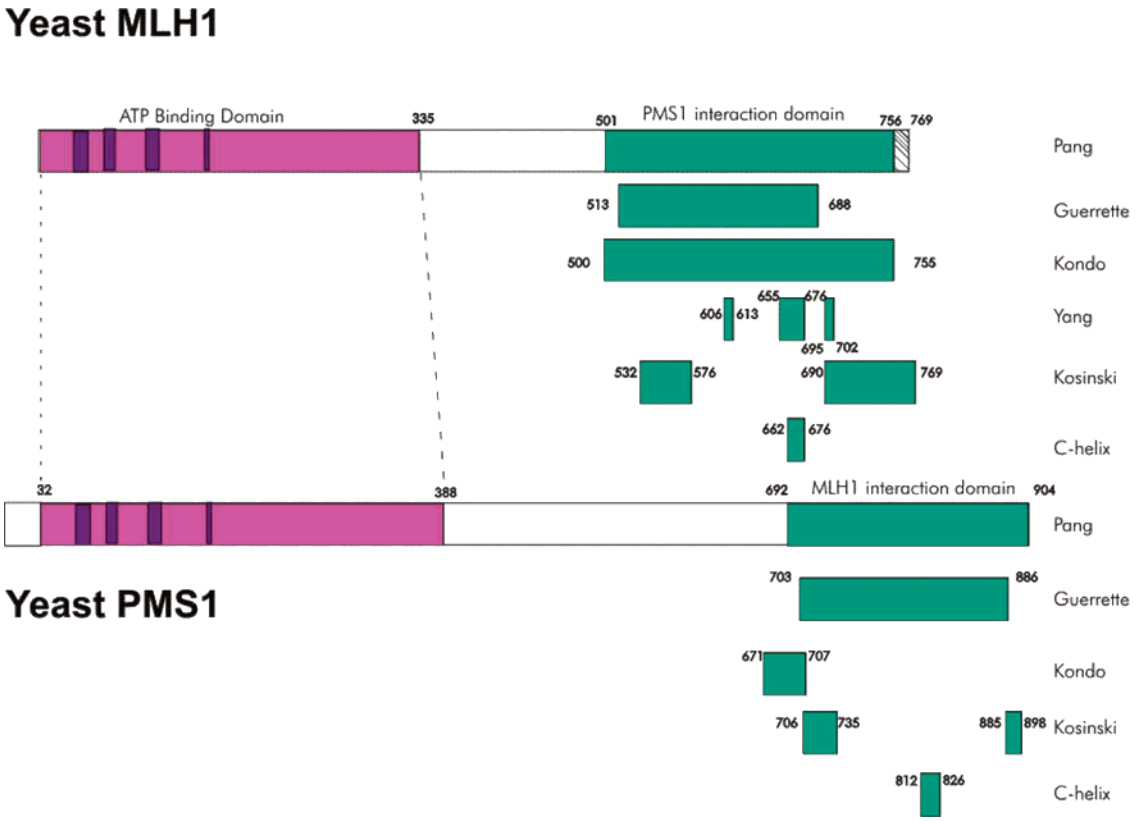

FIGURE 1: Schematic of yMLH1 and yPMS1 organization. The ATP binding domains are shown by homology to MutL. Purple bars represent conserved ATP binding motifs. The MLH1 and PMS 1 interaction domains as determined by Pang et al. (*5*), Guerrette et al. (*7*), Kondo et al. (*6*), Guarné et al. (*9*), and Kosinski et al. (*10*) are colored green.

MATERIALS AND METHODS

Protein Purification. Full-length MLH1 and the MLH1−PMS1 heterodimer were purified as previously described (*11*).

Chemical Modification of Lysine Residues and Proteolysis. Acetylation of lysine residues was performed in a buffer containing 137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \cdot 7H_2O$, and 1.47 mM KH_2PO_4 at pH 7.2. Under these conditions, MLH1 is a monomer at a concentration of 1 μM and forms a stable heterodimer with PMS1. Aliquots containing 1 μM protein in 200 μL were modified with a 10000-, 20000-, 50000-, and 100000-fold molar excess of acetic anhydride (Aldrich, St. Louis, MO) at room temperature for 30 min. MALDI/MS spectra of the resulting modified protein showed no change in the mass of the acetylated protein (within experimental error of the MALDI measurement) with ratios below 50000:1. At higher ratios of acetic anhydride, the mass increased rapidly and the compound began to precipitate. Therefore, the acetylation reactions were conducted at an acetic anhydride:protein ratio of 10000:1 where there was no evidence of protein unfolding. This ratio has been used successfully in previous differential acetylation studies with no evidence of protein unfolding (*12*−*15*). The pH was maintained at 7.0 during the 30 min reaction time by adding sodium hydroxide as needed. The reaction was quenched with 200 μL of 1 M Tris (pH 10.1). A 50 μL aliquot of the modified protein was digested with 1 μg of trypsin (Promega) or Arg-C (Roche) overnight at 37 °C or 1 μg of Glu-C (Roche) or chymotrypsin (Roche) overnight at 25 °C.

Mass Spectrometry. All matrix-assisted laser desorption ionization (MALDI) mass spectra were obtained on an Applied Biosystems (Framingham, MA) Voyager DE STR instrument. MALDI/MS analyses of acetylated proteolysis products were performed in reflector positive mode using an accelerating voltage of 20 kV, a grid voltage of 66%, and a delay time of 350 ns with a mass accuracy of 0.01% or better and a resolution of ≥10000. All analyses were performed using a saturated solution of α-cyano-4-hydroxycinnamic acid in a 45:45:10 (v:v:v) water/ethanol/formic acid mixture as the matrix in a 1:1 ratio with the analyte solution. External calibration was performed using the Applied Biosystems calibration standards.

MALDI/MS/MS analyses were performed using an Applied Biosystems 4700 proteomics analyzer in the positive ion reflector mode. Acetylated peptides were analyzed using 33% saturated α-cyanohydroxycinnamic acid in equal parts water with 0.1% formic acid and acetonitrile with 0.1% formic acid (v:v) as the matrix in a 1:1 ratio with the analyte solution. For MS/MS experiments, the collision energy was 1000 V and argon was used as the collision gas with a recharge threshold set at 1.0×10^{-7} Torr. Calibration was performed externally using the Bruker Daltonics (Billerica, MA) calibration mixture.

For the nano high-performance liquid chromatography (HPLC)/electrospray ionization (ESI)/MS analyses of acetylated peptides, HPLC separations were performed using an Agilent (Palo Alto, CA) 1100 Series nanoCapillary LC device or an Agilent 1100 Series Capillary LC device and a 15 cm × 75 μm (inside diameter) PepMap C18 column (LC

Packings, San Francisco, CA) at a flow rate of 500 nL/min. An 8 μL aliquot of the proteolytic digest was injected. The solvents used with this column were as follows: water and 0.1% formic acid (A) and acetonitrile and 0.1% formic acid (B). The gradient used to elute the peptides began at 5% B for 60 min and then ramped to 95% B over 60 min, followed by a hold at 95% B for 10 min.

All ESI spectra were recorded on a Q-Tof Ultima Global mass spectrometer (Waters/Micromass, Milford, MA), which is equipped with a nanoflow z-spray source. The Q-Tof Ultima Global mass spectrometer was operated at a fwhm resolution of 10 000 in V mode and a mass accuracy of 0.008% or better. The full scan mass spectra were acquired in the positive ion mode over the mass range of *m*/*z* 100−3000 with a source temperature of 80 °C, a capillary voltage of 3.50 kV, a cone voltage of 70 V, a collision energy of 10.0 V, and a scan time of 0.9 s. The collision gas was argon at a gauge pressure of 25 psi. Tandem MS (MS/MS) analyses were performed on the Q-Tof Ultima Global instrument with parameters identical to those used in the MS mode except that the collision energies (selected on the basis of mass and charge state) ranged from 16 to 75 V. Peaks meeting a specified mass/charge state criteria, such as a +3 ion in the mass range of *m*/*z* 50−3000, and with a signal:noise ratio greater than 20 were automatically selected for data-dependent MS/MS acquisition. MS/MS spectra obtained from multiply charged precursors were transformed to singly charged spectra using MaxEnt3 (Waters/Micromass, Altrincham, U.K.). External calibration was performed using renin substrate tetradecapeptide (Sigma, St. Louis, MO).

Sequence Alignments and Homology Models. The crystal structures of the N-terminal domains of *E. coli* MutL (PDB entry 1KBN) (*16*), human PMS2 (PDB entry 1H7S) (*16*), and yeast PMS1 (unpublished 2.1 Å structure; L. C. Pedersen et al., manuscript in preparation) were used as templates to generate a homology model for the N-terminal domain of yeast MLH1 using the Swiss Model first approach mode (*17*). A model for the C-terminal domain of MLH1 was generated in a similar manner using the fold suggested for the *E. coli* MutL C-terminal domain (*9*) and the secondary structural elements predicted using PsiPred (*18*).

RESULTS AND DISCUSSION

Mass spectrometry (MS) has been used to obtain tertiary structural information for a variety of applications, including mapping conformational and discontinuous epitopes (*19*, *20*), studying protein folding and dynamics (*21*), determining protein−protein interactions (*22*−*27*), characterizing protein−nucleic acid interactions (*19*, *28*−*30*), and probing interactions in antibody−antigen complexes (*31*). MS offers advantages over other structural techniques because of the small sample amount required for the analysis, the ability to identify posttranslational modifications, and the capability of working with inhomogeneous samples and with high-molecular weight samples (*32*). Other methods of interrogating protein structure include molecular modeling and chemical surface modification in combination with mass spectrometry (*12*, *33*, *34*).

In our study, we have used the combination of differential chemical surface modification, mass spectrometry, and molecular modeling. In differential chemical surface modification, changes in the extent of chemical modification of proteins when they are part of a complex and when they are in the unbound state provide information about residues on the protein that become shielded in a complex, i.e., when they are part of the interaction surface. This approach compensates for differences in surface reactivity due to, for example, steric effects or the presence of salt bridges. Unfortunately, due to the instability of PMS1 in the absence of MLH1, we could probe the surface reactivity of PMS1 only when it was part of the MLH1−PMS1 heterodimer. However, we could obtain information about PMS1 residues in the heterodimer that are relatively accessible or inaccessible to chemical modification, which does provide structural information.

Acetylation of MLH1 and the MLH1−PMS1 Heterodimer. MLH1 alone and the MLH1−PMS1 heterodimer were acetylated to identify lysine residues on the surface of these proteins. Acetylation was performed at a protein concentration of 1 μM. At this concentration, MLH1 is primarily monomeric on the basis of a dissociation constant of 3.14 μM, while the MLH1−PMS1 heterodimer is primarily dimeric on the basis of its dissociation constant of 87 nM (*9*). The modified proteins were digested with various proteases, and the resulting peptides were analyzed by MALDI/MS, MALDI/MS/MS, and ESI/MS/MS in determining which peptides were acetylated, verifying the peptide sequence, and confirming the site(s) of modification.

A representative of the MS/MS spectra obtained from proteolytic peptides from acetylated MLH1 alone is shown in Figure 2. This is the deconvoluted (+1 charge state) MS/MS spectrum of the ion at *m*/*z* 966.92 [$(M + 2H)^2$], corresponding in mass to a chymotryptic peptide (cleavage at Leu, a secondary cleavage site for chymotrypsin) from MLH1 containing residues 672−686 plus two acetylations. The observed b and y ion series provide sufficient sequence coverage of the peptide to confirm that Lys 675 and Cys 685 of MLH1 are acetylated on the basis of the observation of a +42 Da mass shift, indicating acetylation, for b ions 4−13 and y ion 2.

Figure 3 shows two representative deconvoluted (+1 charge state) ESI tandem mass spectra from the tryptic digest of the acetylated MLH1−PMS1 heterodimer. Panel A shows the MS/MS spectrum of the ion at *m*/*z* 842.03 [$(M + 2H)^2$], which corresponds in mass to tryptic peptide T86−87 from MLH1 (residues 651−665) plus one acetylation. Panel B shows the MS/MS spectrum of the ion at *m*/*z* 811.88 [$(M + 2H)^2$], which corresponds in mass to tryptic peptide T86−87 from PMS1 (residues 679−691) plus one acetylation. The b and y ion series present in each panel provides sufficient sequence coverage of each peptide to confirm that Lys 657 of MLH1 and Lys 680 of PMS1 are, in fact, acetylated by the observation of a mass difference of 170 Da between y_8 and y_9 (for Lys 657) and between b_2 and b_3 (for Lys 680). If this residue were not acetylated, the mass difference would be 128 Da. Additionally, y ions 9−13 and b ions 2−6 (Figure 3A) and b ions 2−4 (Figure 3B) would correspond in mass to a peptide with no modification, and thus, the masses would not incorporate the +42 Da mass shift. The ion series also confirm that Lys 665 of MLH1 and Lys 691 of PMS1 are

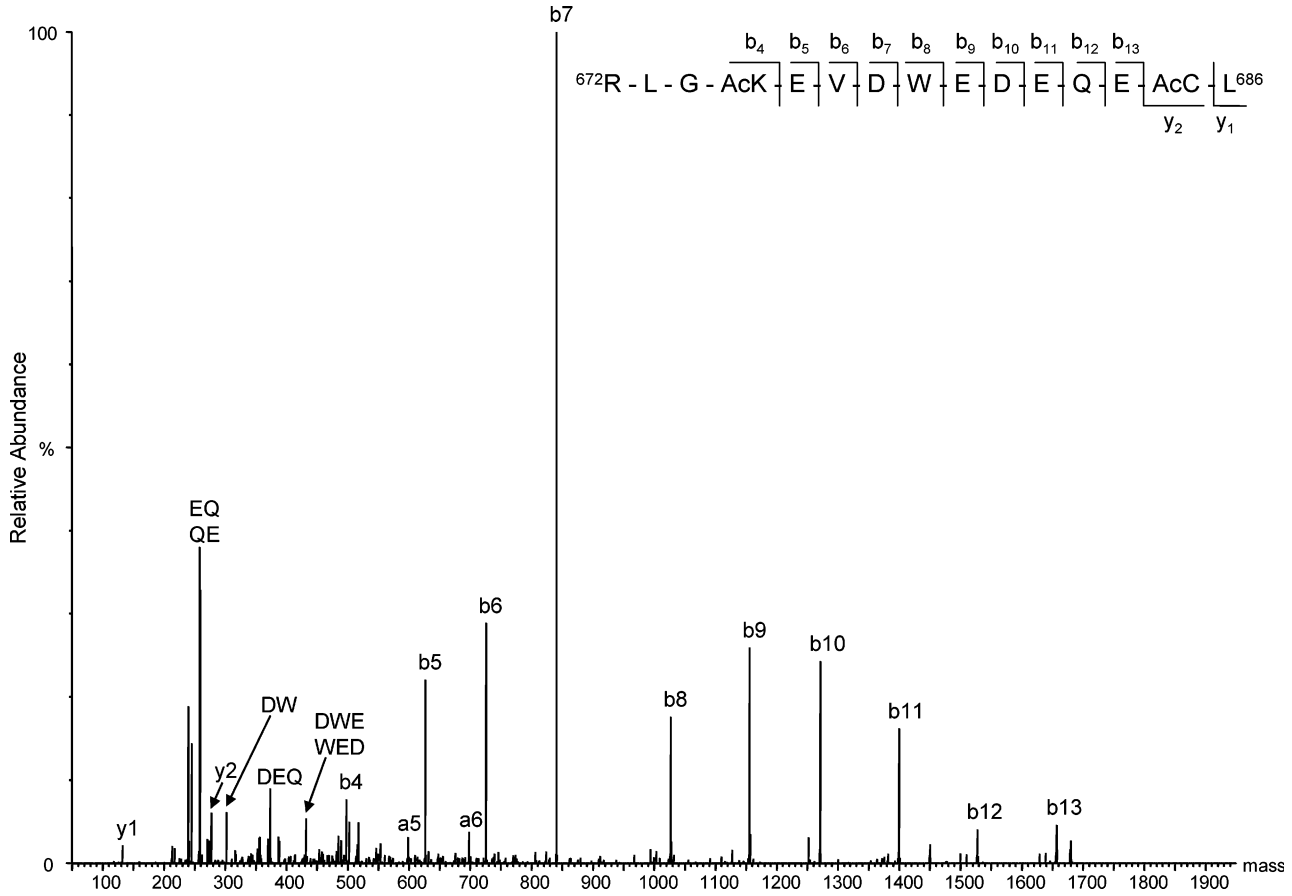

FIGURE 2: Transformed ESI tandem mass spectrum of the ion at *m*/*z* 966.92 [$(M + 2H)^2$] from the chymotryptic digest of acetylated MLH1. The tandem mass spectrum shows this peptide to arise from an anomalous cleavage peptide corresponding to residues 672−686, with acetylated Lys 675 and Cys 685 [$(M + 2H)^2$ calcd 966.94].

unmodified because the mass of the y_1 ion in both spectra is 128 Da.

Figure 4 shows the extracted ion chromatograms (EIC) for Glu-C peptide A89−92 from MLH1 (residues 701−716), with and without modification of Lys 704. For the MLH1 monomer (Figure 4A where both EICs are normalized to 456 counts), the upper EIC, corresponding in mass to acetylated A89−92, exhibits a retention time of 85.0 min. The lower EIC (Figure 4A) which corresponds in mass to unmodified A89−92 exhibits a retention time of 78.1 min. When the abundance of the EICs is normalized to that of the modified A89−92, these data clearly show a significantly higher percent of modified peptide (96%) compared to unmodified peptide (4%) for the monomer of MLH1. The EICs for the same peptide observed from the experiments with the MLH1−PMS1 heterodimer (panel B where both EICs are both normalized to 456 counts) show the same retention times for the peptides with and without modification. However, the abundance of these chromatograms clearly shows a significantly higher percentage of unmodified peptide (97.8%) compared to modified peptide (2.2%) for the heterodimer. These results indicate Lys 704 is modified in the monomer of MLH1, but not in the MLH1−PMS1 heterodimer (Table 1).

The overall results using this approach identified 34 acetylated and 9 nonacetylated lysine residues (66% lysine coverage; Table 1, columns 1 and 2) among 65 total lysine residues in MLH1 alone. Also identified were 40 acetylated and 8 nonacetylated lysines among 78 total lysine residues (62% lysine coverage) from PMS1 in the heterodimer (Table 1, columns 3 and 4). The majority of the detected lysines were confirmed by MS/MS data. A number of lysines were not detected which may be due to the size of the proteolytic peptide, or they may have been suppressed in the complex mixture even using chromatographic separation. Unmodified, but detected, lysine residues (3 of 44 from MLH1 monomer and 8 of 48 from PMS1) indicate that the lysine is not readily accessible to the reagent and/or that reactivity is reduced, possibly due to the presence of a salt bridge. Lysines that are acetylated in the monomeric MLH1 or PMS1 but are observed nonacetylated in the heterodimer are considered to be protected in the heterodimer. There are three residues in MLH1, Lys 665, Lys 675, and Lys 704, which were confirmed by MS/MS to be acetylated in MLH1 alone, but not in the MLH1−PMS1 heterodimer. This protection from modification in the dimer suggests that these three lysine residues are at or near the dimer interface and implies that they may perhaps contribute to dimerization. All three residues are present within the large regions that have been previously implicated in dimerization (*5*−*7*).

Homology Model of the N-Terminal Domain of MLH1. The lysine residues that were acetylated in MLH1 alone (Table 1) are predicted to be on the surface of the protein. To test this prediction in the absence of a structure of MLH1, we generated homology models of the MLH1 N- and C-terminal domains. The homology model for the MLH1

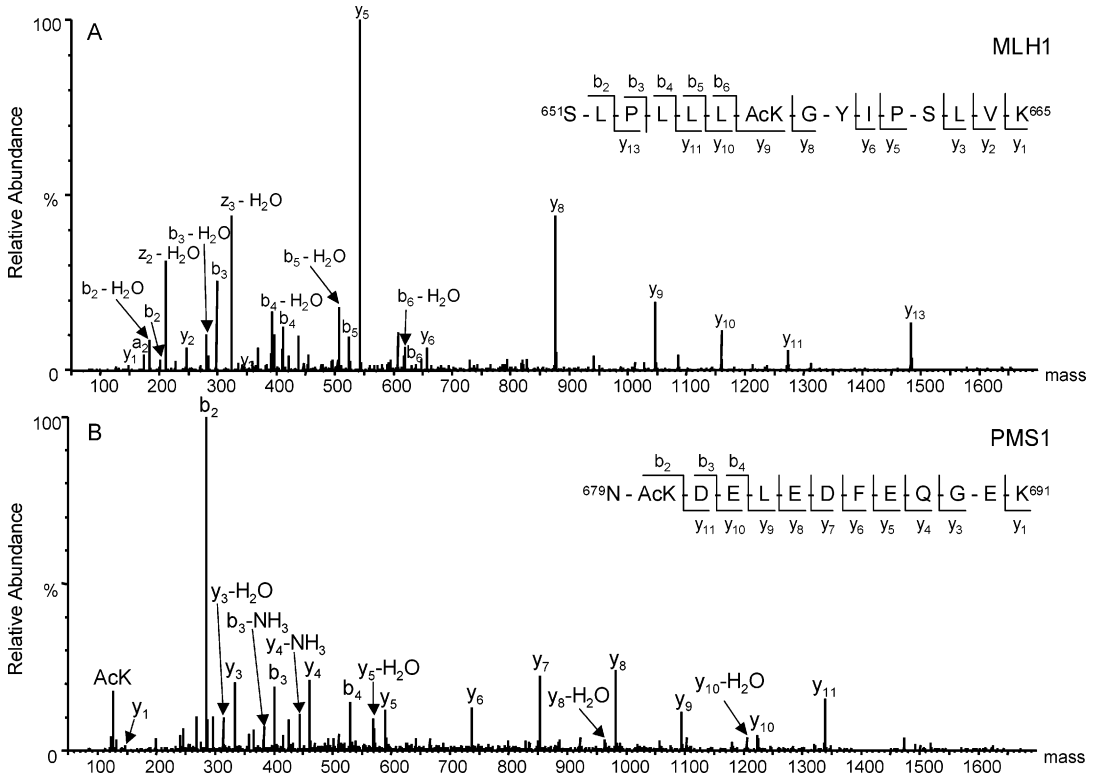

FIGURE 3: Transformed ESI tandem mass spectra from the tryptic digest of the acetylated MLH1−PMS1 heterodimer. (A) MS/MS of the ion at *m*/*z* 842.03 of MLH1, T86−87 (residues 651−665), with acetylated Lys 657 and unmodified Lys 665 [$(M + 2H)^2$ calcd 842.03]. (B) MS/MS of the ion at *m*/*z* 811.88 [$(M + 2H)^2$] of PMS1, T86−87 [residues 679−691; $(M + 2H)^2$ calcd 811.86], with acetylated Lys 680 and unmodified Lys 691. The b and y ion series are labeled and shown on the sequence of each peptide.

N-terminal domain, which contains residues that are well conserved among MutL homologues (*8*), is shown in Figure 5A. In this model, all the lysine residues observed to be acetylated in MLH1 were found on the surface (residues colored blue). Moreover, all the lysine residues observed to be acetylated in the N-terminal domain of yeast PMS1 present in the heterodimer (Table 1) were also found on the surface of the PMS1 N-terminal domain structure (L. C. Pedersen et al., manuscript in preparation), just as predicted. These results are consistent with retention of the native structures under our acetylation conditions.

Model of the C-Terminal Domain of MLH1. Although the C-terminal residues of MutL family members are much less conserved than the N-terminal residues (*8*), it has been suggested that secondary structural elements are conserved (*9*). Therefore, prior to generating a homology model for the C-terminus of yeast MLH1, we performed secondary structure predictions for C-terminal residues of seven eukaryotic MutL proteins, yeast MLH1, MLH2, MLH3, and PMS1 and human MLH1, PMS1, and PMS2. This analysis suggested that most of the β-sheets (1−8) and α-helices (A−G) found in the crystal structure of the *E. coli* MutL C-terminal domain (*9*) were present in the eukaryotic proteins in the same order that was observed in *E. coli* MutL. We used these predicted structural elements to create a refined sequence alignment (Figure 5B) which, together with the fold suggested for the *E. coli* MutL C-terminal domain (*9*), aided development of a molecular model of the C-terminal domain of MLH1 (Figure 5C). As predicted by this model, the lysine residues in MLH1 that are acetylated (blue and pink in Figure 5C) are indeed surface exposed. A few acetylated lysine residues in the N-terminus of MLH1, e.g., 657, 704, and 717, are not shown in the model because they are located in flexible loops and their side chains cannot be accurately modeled.

Relationship of Acetylation Data and Homology Models to the Model of Kosinski et al. On the basis of the crystal structure of the *E. coli* MutL C-terminal domain (*9*), Kosinski et al. (*10*) recently proposed a model for the dimer interface formed by hydrophobic interactions involving residues in MutL corresponding to residues 532−576 and 690−769 in yMLH1 and 706−735 and 885−898 in yPMS1 in our homology model (Figures 1 and 5B). The Kosinski model is consistent with their elegant cross-linking data and with the reduced MMR activity of a mutant in which C-terminal residues 605−615 were deleted. Pang et al. (*5*) also reported loss of MMR activity when the corresponding residues in yMLH1 were deleted. However, they also noted that deletion of these residues did not affect formation of the MLH1−PMS1 dimer. In yeast MLH1, the region of residues 532−576 contains a lysine at position 556 that was not acetylated in either the monomer or the heterodimer and the region of residues 690−769 contains two lysines, K751 and K764, both of which were acetylated in both the monomer and the heterodimer (Table 1). Thus, the results reported by Pang (*5*) and our surface modification results for the yMLH1−PMS1 heterodimer are not consistent with the *E. coli* MutL homodimer model. It is, of course, possible that the heterodimerization domain in yeast differs from the homodimerization domain in *E. coli*. That homodimerization of yeast proteins is less favored is indicated by the much higher K_d for the yeast MLH1−MLH1 homodimer (3.1 μM)

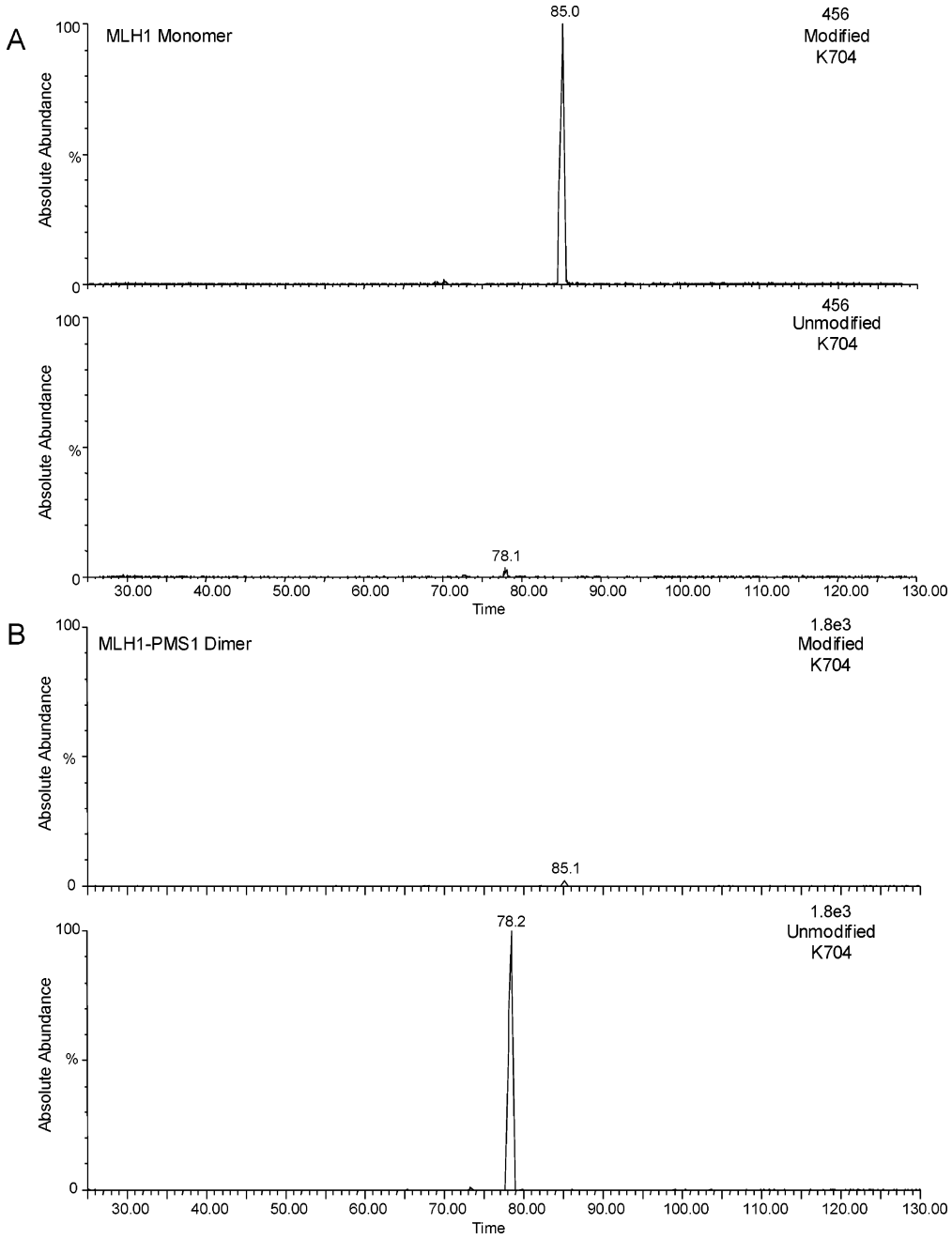

FIGURE 4: Extracted ion chromatograms for Glu-C peptide A89−92 (residues 701−716) shown for the MLH1 monomer (A) with the top chromatogram for the peptide with Lys 704 acetylated and the bottom chromatogram for the peptide with no modification. Panel B shows the MLH1−PMS1 heterodimer single-ion chromatograms for the same peptide with Lys 704 acetylated in the top chromatogram and unmodified in the bottom chromatogram. The minimal time differences are due to manual manipulation of the column at the start of the run.

as compared to that of the yeast MLH1−Pms1 heterodimer (87 nM) (*4*).

Model for the Involvement of C Helices of MLH1 and PMS1 in Heterodimerization. Helix C in *E. coli* MutL is strongly implicated in dimerization (*9*). On that basis, we modeled the interaction of yeast MLH1 with yeast PMS1 in the area of their C helices (Figure 6). In this model, the side chains of Lys 665 and Lys 675 of MLH1, which were protected from acetylation in the heterodimer but not the monomer (Table 1), both face toward PMS1. Lys 665 of MLH1 stacks with His 824 of PMS1, and it may also form a salt bridge or hydrogen bond with Glu 825 of PMS1. Lys 675 of MLH1 could potentially interact with residues located in the loop following helix D, although any such interactions cannot be accurately modeled. Arg 672 of MLH1 may form a salt bridge with Asp 815 of PMS1 and may also hydrogen bond with Glu 818 of PMS1, thereby possibly stabilizing the MLH1−PMS1 interaction.

The yeast PMS1 residues implicated by the model in Figure 6 are all encompassed by the mapping studies of the dimerization interface by Pang et al. (in the yeast MLH1−PMS1 heterodimer) (*5*), Guerrette et al. (in the human MLH1−PMS2 heterodimer) (*7*), and Kondo et al. (in hMLH1 heterodimers but not in interacting partners hMLH3, hPMS1, and hPMS2) (*6*), but not by Kosinski et al. for the *E. coli* MutL homodimer (*10*). Note that Lys 824 of PMS1

Table 1: Acetylation Profile of Lysine Residues in the MLH1−PMS1 Heterodimer[a]

MLH1		PMS1	
N-terminal domain	C-terminal domain	N-terminal domain	C-terminal domain
K6	K344	K14	K408
K15	K352	K17	K411
K30	K370	K19	K453
K49	K391	K27	K498
K54	K393	K60	K499
K67	K398	K79	K504
K81	K408	K106	K531
K84	K427	K111	K537
K115	K434	K114	K551
K117	K446	K120	K555
K131	K467	K139	K614
K137	K471	K150	K618
K142	K472	K153	K622
K172	K473	K165	K636
K185	K475	K172	K649
K192	K480	K190	K652
K193	K489	K194	K656
K204	K496	K197	K659
K219	K504	K202	K667
K232	K515	K218	K669
K244	K516	K227	K678
K253	K520	K229	K680
K254	K556	K230	K691
K286	K583	K244	K699
K311	K613	K271	K703
K325	K615	K277	K704
	K619	K290	K721
	K645	K294	K725
	K648	K309	K739
	K650	K318	K752
	K657	K328	K755
	K665	K335	K781
	K675	K364	K785
	K704	K380	K787
	K717	K396	K799
	K724		K807
	K740		K824
	K751		K838
	K764		K856
			K860
			K861
			K875
			K900

[a] Colors are as follows: brown for acetylated, blue for nonacetylated, green for seen acetylated but not confirmed by MS/MS, purple for unmodified but not confirmed by MS/MS, black for no data, and red, underlined and highlighted in yellow for acetylated in MLH1 alone but unmodified in the MLH1−PMS1 heterodimer.

is acetylated in the heterodimer. In the modeled dimer interface, this side chain is directed at an angle from the dimer interface and is therefore predicted to be surface-exposed. The yeast MLH1 residues implicated by the model in Figure 6 are all encompassed by the mapping studies of the dimerization interface by Pang et al. (*5*), Guerrette et al. (*7*), and Kondo et al. (*6*) but not Kosinski et al. (*10*). The model is further supported by the observation that yeast PMS1 fails to interact with MLH1 containing an R672P substitution, a substitution that results in loss of MMR activity and hereditary nonpolyposis colon cancer in humans.

Implications for Formation of Other MutL Heterodimers. The sequence alignments and secondary structure predictions in Figure 5 and the model in Figure 6 may be relevant to formation of other MutL heterodimers. For example, residues 662−677 of MLH1 may also interact with residues 642−656 of yeast MLH2 or residues 611−626 of yeast MLH3. Similarly, residues 649−664 of human MLH1 may interact

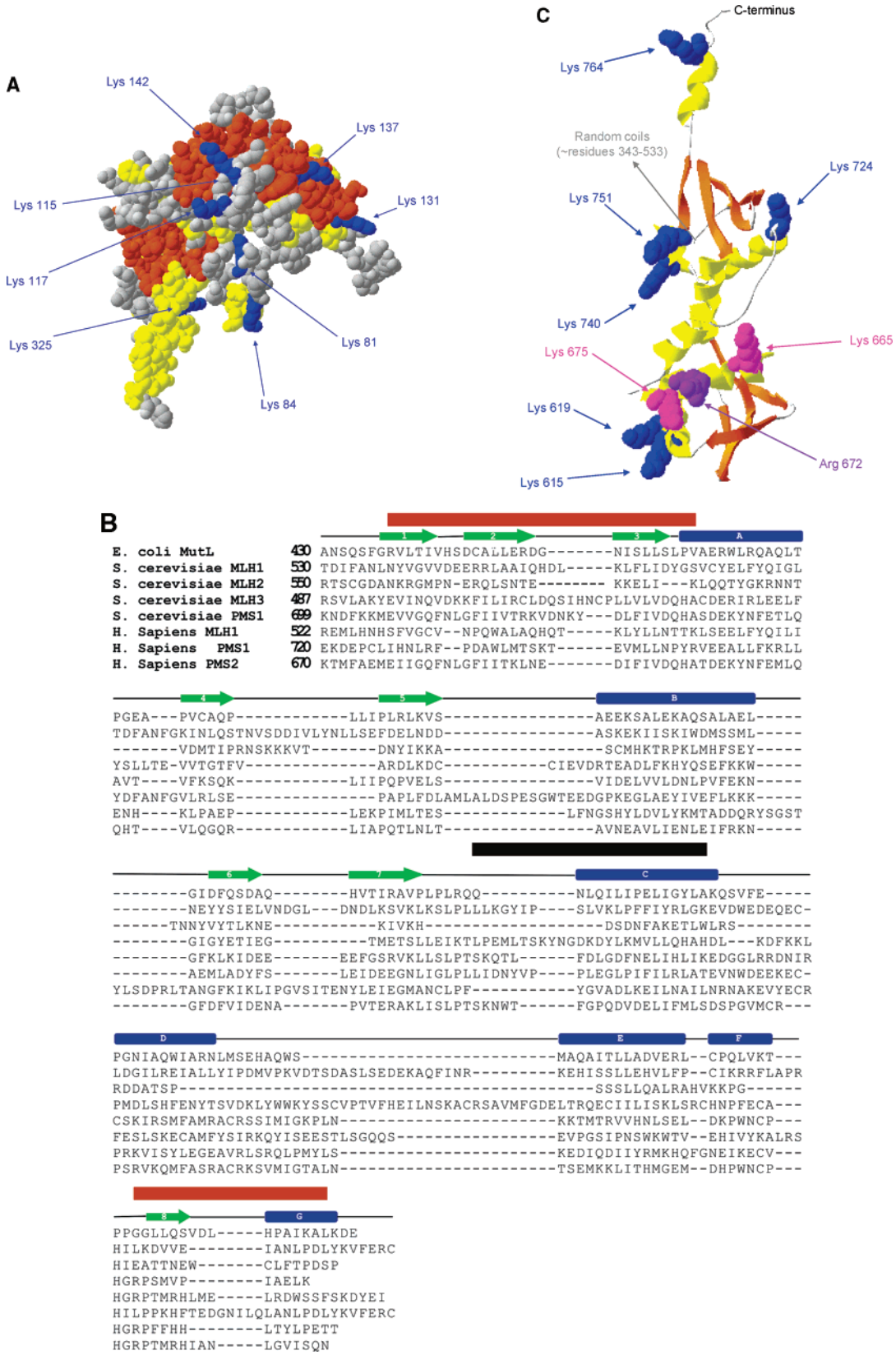

FIGURE 5: (A) MLH1 N-terminal domain homology model with highlighted acetylated Lys residues (blue). Orange denotes β-sheet and yellow α-helix. (B) Secondary structure prediction and alignment for the C-terminal domain of selected members of the MutL family with the structural elements numbered. Boxes with letters represent α-helices; arrows with numbers represent β-sheets. The thick red line denotes the *E. coli* homodimerization domain as proposed by Kosinski et al. (*10*). The thick black line is the predicted heterodimerization domain based on the work of Pang et al. (*5*). (C) Model of the yeast MLH1 C-terminal domain. Lysine residues that were seen to be acetylated have their side chains colored blue. Acetylated residues thought to be involved in DNA binding are colored pink. Arg 672 is colored purple. Orange denotes β-sheet and yellow α-helix. K30, K172, and K311 are on the obversed side of the model. No modification information was obtained for K204 and K249. K67, not observed as modified although on the surface in the model, is in the proximity of Asp119 and may be involved in a salt bridge.

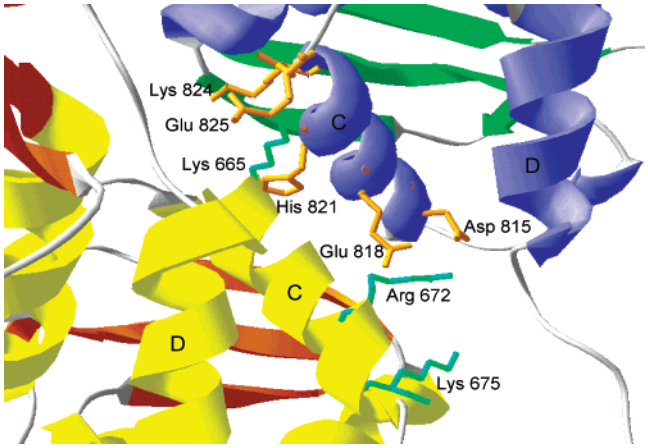

FIGURE 6: Proposed MLH1−PMS1 heterodimer interaction interface. The side chains of residues 665, 672, and 675 of MLH1 (blue) and residues 815, 818, 821, 824, and 825 of PMS1 (orange) are shown. Black letters correspond to the secondary structure prediction: orange (MLH1) and blue (PMS1) for α-helices and yellow (MLH1) and green (PMS1) for *β*-sheets.

with residues 778−793 of PMS2 or residues 845−860 of PMS1. On the basis of sequence alignments (Figure 6), Arg 672 in yeast MLH1 corresponds to Arg 659 in human MLH1, while both human PMS1 and PMS2 have an Asp (residue 849 in PMS1 and residue 782 in PMS2) and a Glu (residue 852 in PMS1 and residue 785 in PMS2) located in positions corresponding to Asp 815 and Glu 818, respectively, in yeast PMS1. It is possible that these residues may contribute to the heterodimeric interactions in the same manner proposed for the yeast MLH1−PMS1 heterodimer.

ACKNOWLEDGMENT

We thank John Fortune, Mercedes Arana, and Lars Pedersen for their assistance with this project and for critical comments on the manuscript.

REFERENCES

1. Kunkel, T. A., and Erie, D. A. (2005) DNA mismatch repair, *Annu. Rev. Biochem. 74*, 681−710.
2. Iyer, R. R., Pluciennik, A., Burdett, V., and Modrich, P. L. (2006) DNA Mismatch Repair: Functions and Mechanisms, *Chem. Rev. 106*, 302−323.
3. Prolla, T. A., Pang, Q. S., Alani, E., Kolodner, R. D., and Liskay, R. M. (1994) MLH1, PMS1, and MSH2 Interactions during the Initiation of DNA Mismatch Repair in Yeast, *Science 265*, 1091−1093.
4. Shcherbakova, P. V., Hall, M. C., Lewis, M. S., Bennett, S. E., Martin, K. J., Bushel, P. R., Afshari, C. A., and Kunkel, T. A. (2001) Inactivation of DNA mismatch repair by increased expression of yeast MLH1, *Mol. Cell. Biol. 21*, 940−951.
5. Pang, Q. S., Prolla, T. A., and Liskay, R. M. (1997) Functional domains of the *Saccharomyces cerevisiae* Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations, *Mol. Cell. Biol. 17*, 4465−4473.
6. Kondo, E., Horii, A., and Fukushige, S. (2001) The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2, *Nucleic Acids Res. 29*, 1695−1702.
7. Guerrette, S., Acharya, S., and Fishel, R. (1999) The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer, *J. Biol. Chem. 274*, 6336−6341.
8. Ban, C., and Yang, W. (1998) Crystal structure and ATPase activity of MutL: Implications for DNA repair and mutagenesis, *Cell 95*, 541−552.
9. Guarne, A., Ramon-Maiques, S., Wolff, E. M., Ghirlando, R., Hu, X. J., Miller, J. H., and Yang, W. (2004) Structure of the MutL C-terminal domain: A model of intact MutL and its roles in mismatch repair, *EMBO J. 23*, 4134−4145.
10. Kosinski, J., Steindorf, I., Bujnicki, J. M., Giron-Monzon, L., and Friedhoff, P. (2005) Analysis of the quaternary structure of the MutL C-terminal domain, *J. Mol. Biol. 351*, 895−909.
11. Hall, M. C., and Kunkel, T. A. (2001) Purification of eukaryotic MutL homologs from *Saccharomyces cerevisiae* using self-cleaving affinity technology, *Protein Expression Purif. 21*, 333−342.
12. Suckau, D., Mak, M., and Przybylski, M. (1992) Protein Surface Topology-Probing By Selective Chemical Modification And Mass-Spectrometric Peptide-Mapping, *Proc. Natl. Acad. Sci. U.S.A. 89*, 5630−5634.
13. Williams, J. G., Tomer, K. B., Hioe, C. E., Zolla-Pazner, S., and Norris, P. J. (2006) The Antigenic Determinants on HIV p24 for CD4+ T Cell Inhibiting Antibodies as Determined by Limited Proteolysis, Chemical Modification, and Mass Spectrometry, *J. Am. Soc. Mass Spectrom. 17*, 1560−1569.
14. Sharp, J. S., Nelson, S., Brown, D., and Tomer, K. B. (2006) Structural characterization of the E2 glycoprotein from Sindbis by lysine biotinylation and LC-MS/MS, *Virology 348*, 216−223.
15. Hochleitner, E. O., Borchers, C., Parker, C., Bienstock, R. J., and Tomer, K. B. (2000) Characterization of a discontinuous epitope of the human immunodeficiency virus (HIV) core protein p24 by epitope excision and differential chemical modification followed by mass spectrometric peptide mapping analysis, *Protein Sci. 9*, 487−496.
16. http://www.ncbi.nlm.nih.gov/Structure/.
17. http://swissmodel.expasy.org//SWISS-MODEL.html.
18. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server, *Bioinformatics 16*, 404−405.
19. Przybylski, M., Kast, J., Glocker, M. O., Durr, E., Bosshard, H. R., Nock, S., and Sprinzl, M. (1995) Mass spectrometric approaches to molecular characterization of protein-nucleic acid interactions, *Toxicol. Lett. 82*−*83*, 567−575.
20. Parker, C. E., Deterding, L. J., Hager-Braun, C., Binley, J. M., Schulke, N., Katinger, H., Moore, J. P., and Tomer, K. B. (2001) Fine definition of the epitope on the gp41 glycoprotein of human immunodeficiency virus type 1 for the neutralizing monoclonal antibody 2F5, *J. Virol. 75*, 10906−10911.
21. Apuy, J. L., Park, Z. Y., Swartz, P. D., Dangott, L. J., Russell, D. H., and Baldwin, T. O. (2001) Pulsed-alkylation mass spectrometry for the study of protein folding and dynamics: Development and application to the study of a folding/unfolding intermediate of bacterial luciferaset, *Biochemistry 40*, 15153−15163.
22. Aurikko, J. P., Ruotolo, B. T., Grossmann, J. G., Moncrieffe, M. C., Stephens, E., Leppanen, V. M., Robinson, C. V., Saarma, M., Bradshaw, R. A., and Blundell, T. L. (2005) Characterization of symmetric complexes of nerve growth factor and the ectodomain of the pan-neurotrophin receptor, p75(NTR), *J. Biol. Chem. 280*, 33453−33460.
23. Farmer, T. B., and Caprioli, R. M. (1998) Determination of protein-protein interactions by matrix-assisted laser desorption/ionization mass spectrometry, *J. Mass Spectrom. 33*, 697−704.
24. Ilag, L. L., Videler, H., McKay, A. R., Sobott, F., Fucini, P., Nierhaus, K. H., and Robinson, C. V. (2005) Heptameric (L12)-(6)/L10 rather than canonical pentameric complexes are found by tandem MS of intact ribosomes from thermophilic bacteria, *Proc. Natl. Acad. Sci. U.S.A. 102*, 8192−8197.
25. McCammon, M. G., and Robinson, C. V. (2004) Structural change in response to ligand binding, *Curr. Opin. Chem. Biol. 8*, 60−65.
26. McCammon, M. G., and Robinson, C. V. (2005) Me, My Cell, and I: The role of the collision cell in the tandem mass spectrometry of Macromolecules, *BioTechniques 39*, 447.
27. Sobott, F., and Robinson, C. V. (2002) Protein complexes gain momentum, *Curr. Opin. Struct. Biol. 12*, 729−734.
28. Deterding, L. J., Kast, J., Przybylski, M., and Tomer, K. B. (2000) Molecular characterization of a tetramolecular complex between dsDNA and a DNA-binding leucine zipper peptide dimer by mass spectrometry, *Bioconjugate Chem. 11*, 335−344.
29. Hanson, C. L., and Robinson, C. V. (2004) Protein-nucleic acid interactions and the expanding role of mass spectrometry, *J. Biol. Chem. 279*, 24907−24910.
30. Steen, H., Petersen, J., Mann, M., and Jensen, O. N. (2001) Mass spectrometric analysis of a UV-cross-linked protein-DNA complex: Tryptophans 54 and 88 of *E. coli* SSB cross-link to DNA, *Protein Sci. 10*, 1989−2001.

31. Tito, M. A., Miller, J., Walker, N., Griffin, K. F., Williamson, E. D., Despeyroux-Hill, D., Titball, R. W., and Robinson, C. V. (2001) Probing molecular interactions in intact antibody: Antigen complexes, an electrospray time-of-flight mass spectrometry approach, *Biophys. J. 81*, 3503−3509.
32. Godovac-Zimmermann, J., and Brown, L. R. (2001) Perspectives for mass spectrometry and functional proteomics, *Mass Spectrom. Rev. 20*, 1−57.
33. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A. 97*, 5802−5806.
34. Steiner, R. F., Albaugh, S., Fenselau, C., Murphy, C., and Vestling, M. (1991) A mass spectrometry method for mapping the interface topography of interacting proteins, illustrated by the melittin-calmodulin system, *Anal. Biochem. 196*, 120−125.

BI061392A