A Protein Pre-Organized to Trap the Nucleotide Moiety of Coenzyme B_{12}: Refined Solution Structure of the B_{12}-Binding Subunit of Glutamate Mutase from *Clostridium tetanomorphum*

Bernd Hoffmann,[a] Martin Tollinger,[a] Robert Konrat,[a] Marja Huhta,[b] E. Neil G. Marsh,[b] and Bernhard Kräutler*[a]

Dedicated to Professor Wolfgang Buckel on the occasion of his 60th birthday

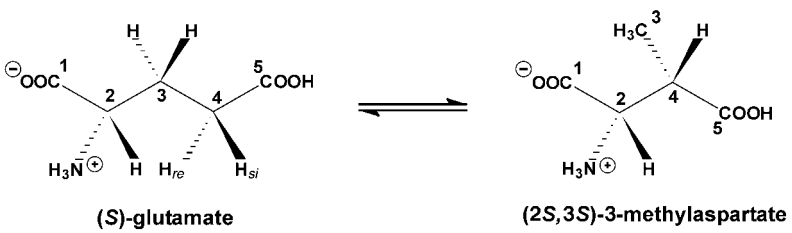
Uniformly ^{13}C,^{15}N-labeled MutS, the coenzyme B_{12}-binding subunit of glutamate mutase from Clostridium tetanomorphum, *was prepared by overexpression from an* Escherichia coli *strain. Multidimensional heteronuclear NMR spectroscopic experiments with aqueous solutions of ^{13}C,^{15}N-labeled MutS provided signal assignments for roughly 90% of the 1025 hydrogen, 651 carbon, and 173 nitrogen atoms and resulted in about 1800 experimental restraints. Based on the information from the NMR experiments, the structure of MutS was calculated, confirming the earlier, less detailed structure obtained with ^{15}N-labeled MutS. The refined analysis allowed a precise determination of the secondary and tertiary structure including several crucial side chain interactions. The structures of (the apoprotein) MutS in solution and of the B_{12}-binding subunit in the crystal of the corresponding homologous holoenzyme from* Clostridium cochlearium *differ only in a section that forms the well-structured helix α1 in the crystal structure and that also comprises the cobalt-coordinating histidine residue. In the apoprotein MutS, this part of the B_{12}-binding subunit is dynamic. The carboxy-terminal end of this section is conformationally flexible and has significant propensity for an α-helical structure ("nascent helix"). This dynamic section in MutS is a decisive element for the binding of the nucleotide moiety of coenzyme B_{12} and appears to be stabilized as a helix (α1) upon trapping of the nucleotide of the B_{12} cofactor.*

KEYWORDS:

coenzyme B_{12} · glutamate mutase · isomerases · NMR spectroscopy · protein structures

Introduction

Glutamate mutase (Glm) from *Clostridium tetanomorphum*[1] and *Clostridium cochlearium*[2] depends upon an adenosyl-cobamide cofactor and catalyzes the reversible carbon skeleton rearrangement of (*S*)-glutamate to (2*S*,3*S*)-3-methylaspartate (Scheme 1). This isomerization is the first step in the fermentation of glutamate to ammonia, CO_2, acetate, butyrate, and H_2 by these bacteria.[3, 4] Glm is composed of two components: E, a homodimer (ε_2, $M_r = 107\,600$ Da) and S, a monomer (σ, $M_r = 14\,700$ Da).[2, 5] The active Glm holoenzyme is an $\varepsilon_2\sigma_2$ heterotetramer[6] which binds two molecules of coenzyme B_{12}.[7, 8] The genes coding for the ε and σ chains have been cloned from both clostridia in *Escherichia coli*. They were designated as *mut* genes in *C. tetanomorphum*[9, 10] and as *glm* genes in *C. cochlearium*.[11] The *mutS*/*mutE* genes code for the subunits MutS and MutE of Glm from *C. tetanomorphum*, while the *glmS*/*glmE* genes code for the corresponding GlmS and GlmE subunits from *C. cochlearium*. MutS and GlmS show 84% identity and were identified as the Glm B_{12}-binding subunits, based on their deduced amino acid sequences[6, 9, 12] (Figure 1). The substrate-binding subunits MutE and GlmE are even 90% identical, but no significant similarity to any other proteins is known. An X-ray crystallographic analysis of reconstituted Glm from



Scheme 1. *Glutamate mutase (Glm) catalyzes a carbon skeleton rearrangement reaction that interconverts (S)-glutamate and (2S,3S)-3-methylaspartate (numbering of carbon centers refers to their position in glutamate).*

[a] *Prof. Dr. B. Kräutler, Dr. B. Hoffmann, Dr. M. Tollinger, Prof. Dr. R. Konrat*
Institute of Organic Chemistry
University of Innsbruck
Innrain 52a, 6020 Innsbruck (Austria)
Fax: (+43) 512-507-2892
E-mail: bernhard.kraeutler@uibk.ac.at

[b] *Dr. M. Huhta, Prof. Dr. E. N. G. Marsh*
Department of Chemistry
University of Michigan
Ann Arbor, MI 48109-1055 (USA)

```
                    10          20          30          40          50
GlmS_Cc     <MEKKTIVLGV  IGSDCHAVGN  KILDHAFTNA  GFNVVNIGVL  SPQELFIKAA
MutS_Ct     <MEKKTIVLGV  IGSDCHAVGN  KILDHSFTNA  GFNVVNIGVL  SSQEDFINAA
MutB_Ps      GRRPRILLAK  MGQDGHDRGQ  KVIATAYADL  GFD VDVGPL  F QTPEETAR 642
MetH_Ec      KTNGKMVIAT  VKGDVHDIGK  NIVGVVLQCN  NYEIVDLGVM  VPAEKILRTA 793

                    60          70          80          90          100
GlmS_Cc     IETKAD AILV  SSLYGQGEID  CKGLRQKCDE  AGLEGILLYV  GGNIVVGKQH
MutS_Ct     IETKAD LICV  SSLYGQGEID  CKGLREKCDE  AGLKGIKLFV  GGNIVVGKQN
MutB_Ps     QAVEADVHVVGV  SSLAGGHLTL  VPALRKELDK  LGRPDILITV  GG VIPEQD 692
MetH_Ec     KEVNAD LIGL  SGLITPSLDE  MVNVAKEMER  QGFT IPLLI  GGATTSKAHT 842

                    110         120         130
GlmS_Cc      WPDVEKRFKD  MGYDRVYAPG  T PPEVGIADL  KKDLNIE> 137
MutS_Ct      WPDVEQRFKA  MGFDRVYPPG  T SPETTIADM  KEVLGVE> 137
MutB_Ps      FD ELR KD   GAVEIYTPG  TVIPESAISLV  KKLRASL 726
MetH_Ec      AVKIEQNYS   GPTVYVQNA  SRTVGVVAALL  SDTQRDD 878
```

Figure 1. *Amino acid sequence alignment of B_{12}-binding proteins having the "B_{12}-binding motif": The deduced amino acid sequences of the B_{12}-binding subunits GlmS of Glm from* C. cochlearium *(GlmS_Cc) and MutS of Glm from* C. tetanomorphum *(MutS_Ct) are aligned with those of the B_{12}-binding domains of Mcm from* Propionibacterium shermanii *(residues 595 – 726; MutB_Ps), and of MetH from* E. coli *(residues 744 – 878; MetH_Ec). Invariant residues are shown in bold type. The numbering is given for the MutS sequence.*

C. cochlearium showed it to contain a B_{12} cofactor at each main interface between the GlmS/GlmE subunits,[8] as assumed earlier.[13]

MutS and GlmS contain the sequence Asp-Xxx-His-Xxx-Xxx-Gly (residues Asp 14 – Gly 19), a fingerprint that represents a conserved "B_{12}-binding" motif.[9] Mutation of either His 16 and Asp 14 in MutS significantly weakens coenzyme binding and slows catalysis by about 1000-fold.[7] Other B_{12}-dependent enzymes that contain the "B_{12}-binding" motif include the B_{12}-binding domain of cobalamin-dependent methionine synthase (MetH) from *E. coli* and the B_{12}-binding domains of methylmalonyl-CoA mutase (Mcm)[9] and 2-methyleneglutarate mutase[14] (Figure 1).

The crystal structures of B_{12}-dependent enzymes uncovered the role of the residues of the "B_{12}-binding" motif for cofactor binding and catalysis. Crystallographic work on the methylcobalamin-binding domain metH of methionine synthase from *E. coli* provided the first detailed structure of a B_{12}-dependent enzyme[15, 16] and revealed that the conserved histidine residue of the "B_{12}-binding" motif replaces the 5,6-dimethylbenzimidazole (DMB) moiety as the ligand to cobalt (Figure 2). This replacement results in a "base-off/His-on" form of the methyl-Co^{III}-corrin **2** in methionine synthase. The displaced nucleotide tail is buried in a hydrophobic cleft and serves to anchor the coenzyme to the protein. The cobalt-coordinated histidine residue is involved in an H-bonding network with an aspartate and a serine residue. This set of three residues was named the His-Asp-Ser regulatory triad and has been suggested to control the coordination properties of the cobalt corrin during the catalytic cycle in methionine synthase.[15, 16]

More recently, the crystal structures of recombinant Glm from *C. cochlearium*, reconstituted with methylcobalamin (**2**) and with cyanocobalamin (**3**) (Figure 2), were solved.[8] The binding of the B_{12} cofactors to GlmS was also observed in a "base-off/His-on" constitution. The DMB base is deeply buried between the β sheet (strand β3 and β4) and the two α helices α1 and α5 of GlmS. The DMB base is replaced by the imidazole of His 16, which coordinates to the cobalt center through N^{ε}, while N^{δ} is hydrogen-bridged to Asp 14. The first available X-ray crystal structure of an adenosylcobalamin-dependent enzyme was that of Mcm from *Propionibacterium shermanii*,[17, 18] which also revealed striking similarities in the mode of binding of the cofactor to that of metH (and of Glm).

Editorial Advisory Board Member:[*]
Bernhard Kräutler,

born in Dornbirn (Austria), studied chemistry at the ETH in Zürich, where he also carried out his dissertation under Prof. Albert Eschenmoser. After postdoctoral stays with Profs. A. J. Bard (Austin, Texas) and N. J. Turro (Columbia University), he returned to the ETH and obtained his habilitation in organic chemistry in 1985. From 1991 he has been Full Professor of Organic Chemistry at the University of Innsbruck. He is a corresponding member of the Austrian Academy of Science. His main research interests—bioorganic chemistry, porphyrinoid compounds, coenzyme B_{12}, chlorophyll, fullerenes, structure – function relationships of biomacromolecules, and supramolecular chemistry—are documented in about 130 publications.

[*] *Members of the Editorial Advisory Board will be introduced to the readers with their first manuscript.*

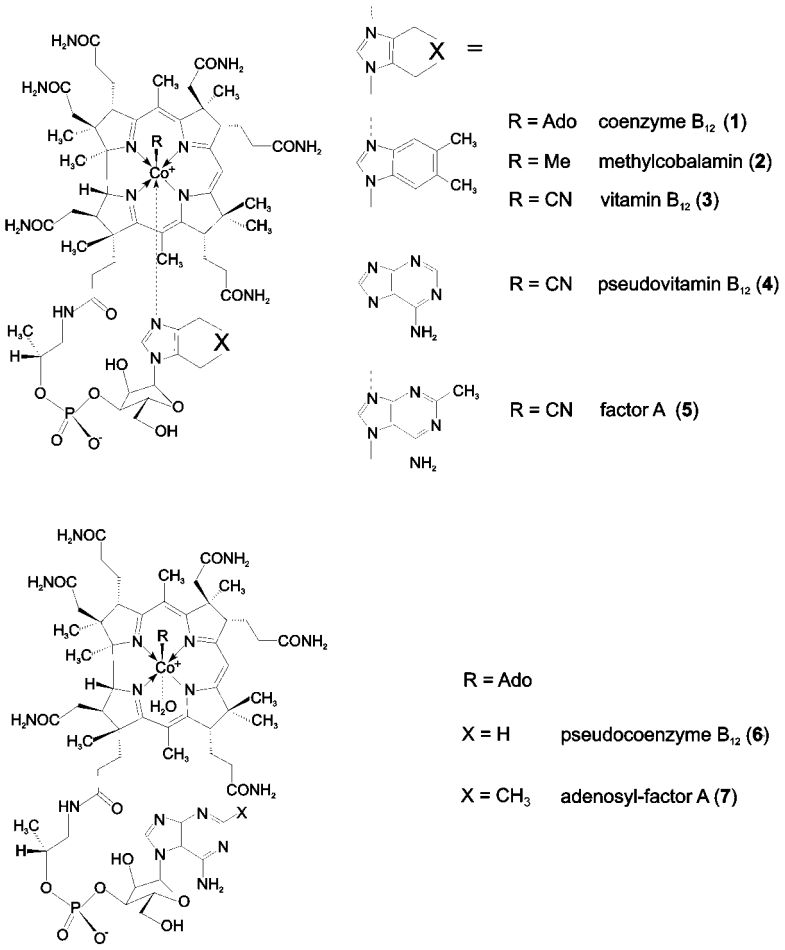


Figure 2. *Structural formulae of B_{12} derivatives. Top ("base-on" cobamides): coenzyme B_{12} (**1**, Ado = 5′-deoxy-5′-adenosyl), methylcobalamin (**2**, Me = methyl), vitamin B_{12} (**3**), pseudovitamin B_{12} (**4**), factor A (**5**); bottom ("base-off" cobamides): pseudocoenzyme B_{12} (**6**) and adenosyl-factor A (5′-deoxy-5′-adenosyl-2-methyladeninyl-cobamide; **7**).*

In all of these holoproteins (metH, Mcm, and GlmS), the B_{12}-binding segments or subunits fold as an α/β domain, in which five (six) α helices encase a β-sheet, that consists of five parallel strands. This domain has a topology similar to that of the nucleotide-binding Rossmann folds.[19] The B_{12} cofactors are bound at the C-terminal end of the β sheet and at the main interface between the B_{12}-binding and substrate-activating domains.[8, 14, 17, 18] Interestingly, glutamate mutase from *C. tetanomorphum* was found not to use coenzyme B_{12} (**1**) as its native cofactor, but its analogue pseudocoenzyme B_{12} (**6**, 5′-deoxy-5′-adenosyl-adeninylcobamide) (Figure 2).[20] More recently, also the main corrinoids from *C. cochlearium* were analyzed (as their Co_β-cyano derivatives) as the well-known naturally occurring adeninyl-cobamides pseudovitamin B_{12} (**4**) and factor A (**5**) (Figure 2).[21] These findings are quite intriguing since the deduced B_{12} coenzyme forms in Glm are indicated to already predominantly exist as "base-off" forms in aqueous solutions.[20, 22]

Binding of B_{12} cofactors and their activation by B_{12}-dependent enzymes are a subject of ongoing studies in our laboratories.[13, 23] Recently, the solution structures of the ^{15}N-labeled B_{12}-free apo forms of the B_{12}-binding subunits of Glm were determined by NMR spectroscopy.[24, 25] These studies showed that both B_{12}-binding subunits were folded globular proteins which resembled the B_{12}-binding domains of the holoenzyme crystal structures of Glm, Mcm, and metH. However, in both MutS and GlmS, one contiguous region (comprising residues Ser 13 – Phe 27 in MutS, residues Gly 9 – Phe 27 in GlmS) was found to be conformationally disordered (in the absence of the B_{12} coenzyme). This sequence contains the conserved B_{12}-binding Asp-Xxx-His-Xxx-Xxx-Gly motif and corresponds to a structured loop and to helix α1 in the holoenzyme crystal structures. Its structural and dynamic properties are crucial for B_{12} binding, as it contains the cobalt-coordinating histidine residue and also forms one side of the B_{12}-nucleotide-binding cleft.

The dynamic segment of MutS, corresponding to helix α1, was found to exhibit a partially formed α-helical structure and to behave as a "nascent" helix,[24] with considerable propensity for an α-helical conformation. On the other hand, ^{15}N NMR relaxation data pointed to rapid conformational processes: The internal mobility therefore was attributed to the interconversion between a holoprotein-like α-helical conformation and random-coil states. These results suggested that helix α1 would be stabilized only upon binding of the corrinoid coenzyme to MutS and GlmS, for example, by hydrophobic packing against the nucleotide tail of the coenzyme, thereby completing the hydrophobic cleft.[24, 25] MutS and GlmS thus furnish a dynamically preformed binding site for the nucleotide moiety of cobamide cofactors. Specific binding of the nucleotide moiety of coenzyme B_{12} (**1**) to MutS indeed was observed in NMR spectroscopic studies and occurred with the suggested α-helical structuring of the "nascent" helix α1, thereby also completely structuring this B_{12}-binding subunit.[26]

To refine the information on the three-dimensional solution structure and the dynamic properties of the corrinoid-free subunit MutS of Glm from *C. tetanomorphum* by using heteronuclear NMR spectroscopy, we prepared uniformly doubly ^{13}C- and ^{15}N-labeled MutS. On the basis of an increased number of NMR spectroscopically derived structure restraints, the resulting MutS structures could be considerably improved. The doubly labeled MutS sample provided additional NMR-derived backbone dihedral angle restraints, crucial for the description of the structure and dynamics of the "nascent" helix α1. Furthermore, important H-bonding and other side chain interactions could be investigated, that connect elements of secondary structure of MutS. From such refined solution structures, detailed insights were expected into the structure and dynamics of the B_{12}-binding protein MutS and into the crucial properties relevant for recognition and binding of B_{12} cofactors in their remarkable "base-off/His-on" form.

Results

Signal assignment

Buffered aqueous solutions of uniformly ^{13}C,^{15}N-labeled MutS, obtained from overexpression in an *E. coli* strain, were subjected to heteronuclear NMR spectroscopic analyses. The signals of the amide groups as well as the C^{α} and C^{β} signals were assigned sequentially by combining the information contained in the HNCA, HNCACB, and CBCA(CO)NH triple-resonance spectra. Assignment of most of the side chain carbon and hydrogen signals was achieved with the HCCH-TOCSY experiment. Hydrogen signals of aromatic side chains (H^{δ} and H^{ε}) were assigned by employing two-dimensional (2D) (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE experiments. Assignment of carbonyl carbons was obtained by inspection of the three-dimensional (3D) HNCO spectrum. For the 133 non-proline residues of MutS, signals for 121 backbone amide functions were observed in the ^{1}H,^{15}N HSQC spectrum under the experimental conditions employed; for twelve backbone amide NH groups signals were not seen, presumably due to fast exchange with bulk water at pH 6.0 and/or rapid conformational exchange (for Met 1 and Glu 2, and residues Cys 15, His 16, Val 39, Leu 40, Ser 42, Leu 63, Tyr 64, Gly 65, Asn 93, and Val 95, all located in the loops at the C-terminal ends of the β strands). For all except four residues (Met 1, Cys 15, Pro 118, Pro 119) C^{α}/H^{α} and side chain signals could be assigned in the HCCH-TOCSY experiment. In addition, a number of side chain NH functionalities could be assigned from the ^{1}H,^{15}N TOCSY-HSQC and HNCACB spectra. These are the ε-NH groups of all three Arg residues, the δ-NH_2 groups of all seven Asn residues, the ε-NH_2 groups of all four Gln residues, and the indolic NH function of the only tryptophan residue, Trp 101. Approximately half of the signals for valine and leucine methyl groups were nondegenerative and displayed differential cross peak patterns in the ^{1}H,^{13}C NOESY-HSQC spectrum. This allowed for the stereospecific assignment of the signals of 7 valine and 5 leucine methyl groups (out of a total of 14 valine and 9 leucine residues in the MutS sequence). The signals of the remaining valine and leucine methyl protons and carbon atoms could not be assigned stereospecifically, as they either were isochronous, indicating conformational averaging, or displayed only weak and stereochemically ambiguous nuclear Overhauser effects (NOEs).

Secondary structure of MutS

Elements of regular secondary structure were previously identified for ^{15}N-labeled MutS on the basis of their specific cross peak patterns, secondary ΔH^{α} shifts (i.e., the difference in chemical shift between experimental and random-coil H^{α} values), $^{3}J(H^{N}H^{\alpha})$ coupling constants and amide proton attenuation factors.[24] The complete set of $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ secondary chemical shifts now obtained for doubly labeled MutS provide a solid qualitative basis for the description of the secondary structure[27] and are shown in Figure 3 as a function of the MutS sequence.

The nearly complete assignment of the ^{1}H, ^{13}C, and ^{15}N signals of ^{13}C,^{15}N-labeled MutS provided an excellent basis for the application of TALOS[28] for a better description of the secondary structure of MutS. TALOS uses sequence information and secondary ΔH^{α}, ΔC^{α}, ΔC^{β}, ΔCO, and ΔN^{H} chemical shifts in order to generate predictions for both protein backbone torsion angles, Φ and Ψ, of each residue. The majority of Φ and Ψ predictions for MutS were classified by TALOS as "good" (unambiguous). Most of these Φ and Ψ values are clearly consistent with those typically found for residues located in α helices and β sheets and fall within the elements of regular secondary structure as determined for ^{15}N-labeled MutS.[24] Almost all predictions classified as "ambiguous, but not necessarily wrong" [a(bnnw)] fall within the loop regions of MutS.

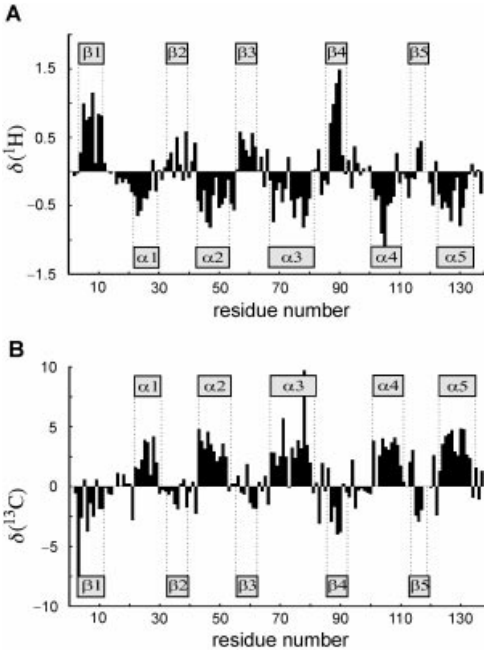


Figure 3. *ΔH^{α} and ΔC^{α} secondary chemical shifts of doubly labeled MutS as a function of residue number (ΔH^{α} and ΔC^{α} are the differences in chemical shifts between experimental and random-coil values for H^{α} and C^{α}, respectively). Elements of regular secondary structures are indicated by labeled bars.*

Except for Thr 28 the ΔH^{α} secondary shifts of sequentially assigned residues Ile 22 to Asn 29, which are located in the "nascent" helix α1, are clearly negative, whereas the ΔC^{α} secondary shifts of the same stretch of residues exhibits positive values. These data indicate this segment to have a pronounced propensity for an α-helical conformation. The TALOS predictions for residues Leu 23 – Asn 29 have mutually consistent values of Φ and Ψ (Φ: $-63 \pm 3^{\circ}$, Ψ: $-39 \pm 6^{\circ}$) and fall into the α-helical region of the Ramachandran plot.[19] TALOS classified the Φ and Ψ predictions for residues Leu 23 – Asn 29 as "good", and they were used as angular restraints in one set of subsequent structure calculations.

Inspection of secondary shifts and NOE patterns as well as the TALOS output revealed several interesting local features of the secondary structure of MutS in solution (Figure 4). Overall, the predicted backbone structure and its torsion angles in segments with defined secondary structure of MutS reproduce remarkably well those in the crystal structure of the holoenzyme Glm from

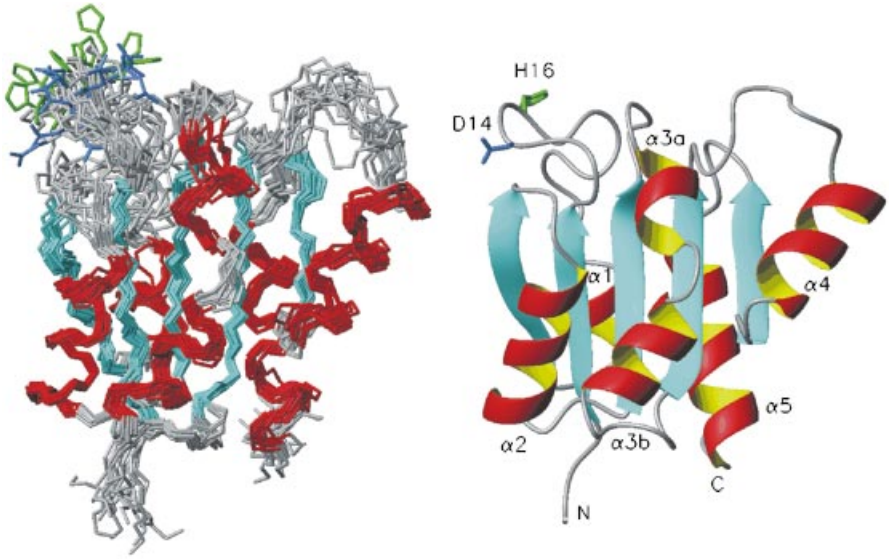


Figure 4. *Solution structure(s) of MutS from models obtained with neglect of TALOS predictions[28] of the Φ and Ψ torsion angles of the "nascent" helix α1 (Leu 23 – Ala 30), represented in a view emphasizing the conformationally disrupted helix α3. Left: Backbone traces (N, C^{α}, carbonyl C) of the best 15 NMR-derived structures of MutS. Right: Ribbon model of one representative structure from the family of 15 structures. Elements of regular secondary structure are labeled according to the scheme used earlier for the description of the crystal structure of the B_{12}-binding domain of MetH.[15] The side chains of Asp 14 and His 16 are highlighted as green stick drawings. (The figures were generated with the program MOLMOL.[63])*

C. cochlearium,[8] including the disruption of helix α3 and a conformational irregularity of the β2 backbone.

Conformational irregularity in β2: Analysis of $H^N - H^N$ and $H^N - H^{\alpha}$ NOEs detected for the backbone amide protons of the MutS residue pair Ile 37 – Gly 38 located in β strand β2 suggest the presence of a main chain conformational "disturbance" of these residues. According to these NOEs, the NH bond vectors of Ile 37 and Gly 38 point to the same side of the β strand (toward β strand β1) contrasting the situation in regular β strands, where the NH bond vectors of successive residues point in opposite directions. The presence of a conformational "irregularity" in the backbone of β2 is corroborated by ΔH^{α} and ΔC^{α} secondary shifts of Ile 37 and Gly 38 which clearly are not in accordance with a β-sheet structure, and the TALOS Φ and Ψ predictions for Ile 37, Gly 38, and Val 39 in β2 are classified as a(bnnw). For Ile 37, the calculated backbone torsion angles ($\Psi = 28 \pm 7^{\circ}$, $\Phi = -73.5 \pm 5^{\circ}$) fall outside of the range of the values for β strands, but are close to those of α helices.[19]

Disruption of helix α3: The ΔH^{α} and ΔC^{α} secondary shifts of Gly 73 (given in Figure 3), located in the middle of helix α3, were not consistent with values expected for a single α-helical stretch extending over the residues 66 – 81. A disruption of helix α3 (see Figure 5) is supported by the backbone dihedral angles Φ and Ψ of this segment classified as a(bnnw) by TALOS for Lys 72 and Gly 73. In addition, the sequential and medium-range NOE patterns observed in the $^{1}H,^{15}N$ NOESY-HSQC spectrum for residues Glu 68 – Gln 76 deviate from a pattern typical for α helices,[29] but support also the conformational disruption of helix α3 near Lys 72 and Gly 73. A total of 65 interresidue NOE restraints was used for the four residues Cys 71 – Leu 74, and the non-helical conformation of this segment is well defined. In contrast, only 20 interresidue NOE restraints were employed for the same segment in the calculations for ^{15}N-labeled MutS.[24] Seven of the 65 interresidue NOEs observed with $^{13}C,^{15}N$-labeled MutS were assigned to through-space interactions between side chain amide protons of Gln 43 and H^N, H^{α} as well as side chain protons of residues Cys71–Leu 74 (Figure 5). These NOEs suggest a long-range interaction between the amide function of Gln 43 and the short loop linking the helices α3a and α3b. The solution structure obtained for $^{13}C,^{15}N$-labeled MutS showed the side chain of Gln 43 (at the N terminus of helix α2) to be pointing towards helix α3 (Figure 5). The side chain amide function of Gln 43 appears to be partially inserted into the gap, which is formed by a 360° turn of helix α3, thereby disturbing the helical conformation of the α3 main chain in the range of residues

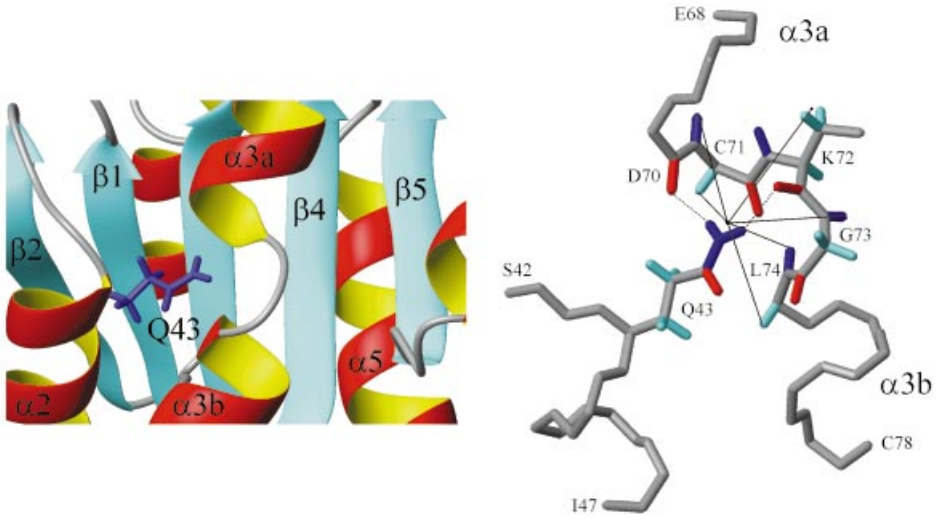


Figure 5. *Disruption of helix α3 in the MutS solution structure. Left: Ribbon drawing close-up view of a representative MutS solution structure showing the disrupted helix α3 and the side chain of Gln 43 (blue). Right: Experimentally observed NOEs between the side chain amide hydrogen atoms of Gln 43 and hydrogen atoms of residues Cys 71 – Leu 74 (solid lines). Potential hydrogen bonds involving the side chain amide hydrogen atoms of Gln 43 are shown as dotted lines. (The figures were generated with the program MOLMOL.[63])*

Cys 71 – Leu 74. Consequently, the backbone amide functions of residues Gly 76 – Lys 72 are prevented from forming hydrogen bonds with the carbonyl oxygen atoms of their $i-4$ predecessors. Instead, the side chain amide function of Gln 43 is in a position where both amide hydrogen atoms could be hydrogen-bonded to the backbone carbonyl oxygen atoms of Asp 70 and Lys 72. In the holoenzyme crystal structure of Glm from *C. cochlearium*[8] disruption of helix α3 and the hydrogen bonds between the side chain amide function of Gln 43 and the backbone carbonyl oxygen atoms of residues Asp 70 and Lys 72 are clearly discernible, in a mode that is highly similar to the situation in the MutS solution structure (Figures 5 and 6 D, F).

The "nascent" helix α1: ^{15}N-relaxation data of MutS[24] indicated the residues Ser 13 – Gly 19 to be involved in significant mobility on a picosecond to nanosecond time scale, residues between Ile 22 and Thr 28 to exchange on a time scale of 10 – 100 μs. The latter internal mobility was attributed to the interconversion between random-coil states and an α-helical conformation. In ^{13}C,^{15}N-labeled MutS, all residues between the C-terminal end of β1 and the N-terminal end of β2 gave rise to detectable amide signals, except Cys 15 and His 16. Only a very limited number of interresidue NOEs were observed along this segment. The NOE pattern found for residues downstream from Ala 30 did not correspond to an α-helical conformation. Medium- and long-range NOEs such as $H^N(i)-H^N(i-2)$, $H^N(i)-H^\alpha(i-3)$, and $H^N(i)-H^\alpha(i-4)$ NOEs, which were typically found for α helices,[29] were only detected for the residues between Ser 26 and Ala 30, the residues closest to the C terminus of the "nascent" helix. On the basis of this limited number of NOE restraints, a short helix α1 was formed in the course of structure calculations (second run of structure calculations); the MutS residues between Gly 12 and His 25 occurred in random-coil conformations (Figures 4, 6 C). ΔH^α and ΔC^α secondary shifts of MutS residues Ile 22 – Asn 29 are clearly consistent with an α-helical conformation (Figure 3). The (time-averaged) population of an α-helical conformation of the segment between residues Ile 22 – Asn 29 must therefore be considerable. Consistent with the observed secondary chemical shifts, the TALOS database system predicts nonambiguous Φ and Ψ backbone torsion angle restraints which indicate residues Ile 22 – Ala 30 to form an α helix (see Figure 6 D).

Structure calculations

Two families of 15 low-energy structures were calculated by using the procedure described above employing a final data input set with a total of 1792 and 1776 NMR-derived restraints, respectively, giving an average of approximately 13 restraints per residue. Thus, the total number of NMR restraints was almost doubled with respect to the previous structure determination of singly ^{15}N-labeled MutS.[24] An increase by a factor of 3 was also obtained for the class of long-range NOEs ($|i-j|>4$), which are a decisive experimental support for defining the tertiary structure of a protein. By use of the TALOS program[28] the number of backbone torsion angle restraints was tripled. A summary of the restraints and of the structural statistics for the final families of structures is given in Table 1.

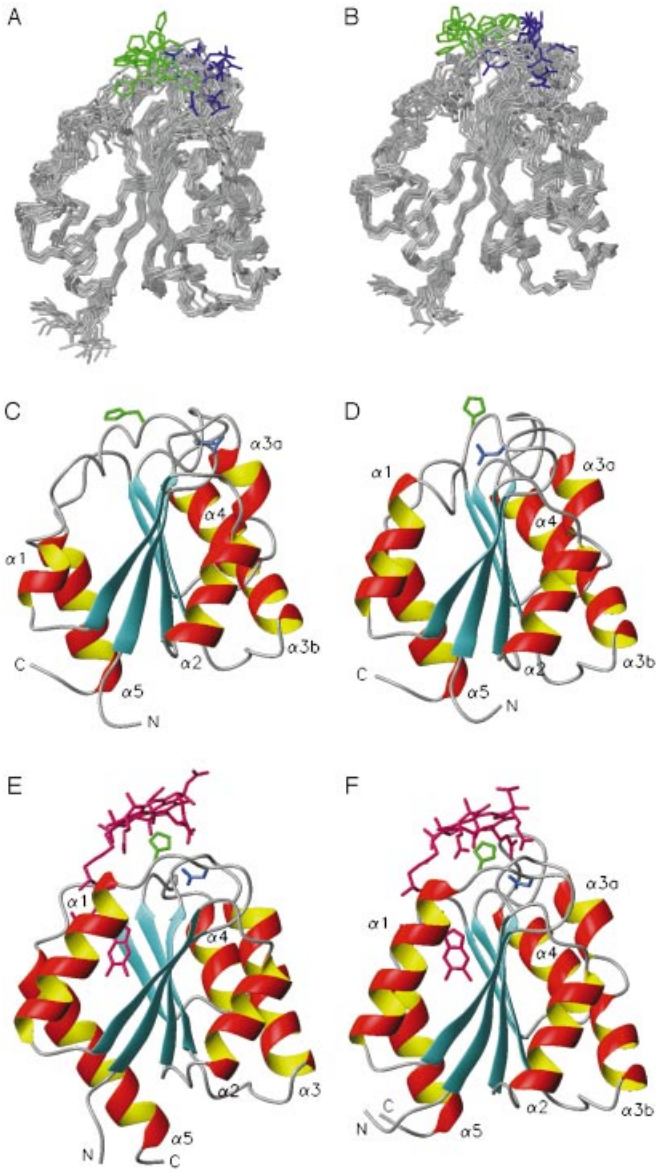


Figure 6. *Top: Backbone traces (N, C$^\alpha$, carbonyl C) of the 15 best NMR-derived structures of MutS. A and B are derived from models obtained with neglect or inclusion of TALOS predictions[28] of the Φ and Ψ torsion angles of the "nascent" helix α1 (Leu 23 – Ala 30), respectively; the side chains of the cobalt-coordinating histidine residue (His 16) and of Asp 14 are depicted in green or dark blue, respectively. Center and bottom: Ribbon models of the B_{12}-binding domains of several coenzyme B_{12}-dependent enzymes. C and D are representative solution structures of MutS from* C. tetanomorphum. *C represents the $\alpha_4\beta_5$ fold of MutS which equilibrates with its $\alpha_5\beta_5$ fold, shown in D, both obtained from the analysis of ^{13}C,^{15}N-labeled MutS. E and F are sections of the crystal structures of the B_{12}-binding domain (mutB, residues 595 – 726) of Mcm of* P. shermanii*[17] and of the B_{12}-binding subunit GlmS of Glm from* C. cochlearium.*[8] Side chains of the cobalt-coordinating histidines His 610 (mutB), His 16 (MutS, GlmS), of residues Asp 608 (mutB), Asp 14 (MutS, GlmS) and the bound B_{12} cofactor of mutB and GlmS are depicted as green, blue, and red stick drawings, respectively. (All figures were generated using the program MOLMOL.[63])*

Tertiary structure of MutS

The tertiary structure of MutS, revealed by analysis of the ^{13}C,^{15}N-labeled protein (Figures 4, 6 A, B) confirms the one determined for ^{15}N-labeled MutS and GlmS.[24, 25] The core of the molecule

Table 1. *Structural statistics for the final MutS structures.*

	15 MutS structures	
	first set ($\alpha_5\beta_5$ fold)[a]	second set ($\alpha_4\beta_5$ fold)[a]
Experimental restraints		
total	1792 (991)[b]	1776
intraresidue	403 (335)[b]	403
interresidue, sequential ($\|i-j\|=1$)	476 (299)[b]	476
interresidue, medium-range ($2\leq\|i-j\|\leq 4$)	263 (119)[b]	263
interresidue, long-range ($\|i-j\|>4$)	411 (125)[b]	411
dihedral angle restraints	184 (57)[b]	168
hydrogen bonds	55 (56)[b]	55
Average rmsd from experimental restraints		
distance restraints [Å $\times 10^{-2}$]	7.0 ± 0.3	7.3 ± 0.3
dihedral angle restraints [°]	0.25 ± 0.03	0.25 ± 0.02
Average rmsd from idealized covalent geometry		
bonds [Å $\times 10^{-2}$]	1.0 ± 0.01	1.0 ± 0.01
angles [°]	2.63 ± 0.02	2.67 ± 0.03
impropers [°]	2.06 ± 0.10	2.12 ± 0.21
X-PLOR energies[c]		
average E_{L-J} [kcal mol^{-1}]	-1397 ± 20	-1377 ± 22
PROCHECK statistics[d]		
residues in allowed regions of Ramachandran plot [%]	100	100
Atomic rmsd [Å][e]		
residues 4 – 10, 23/26 – 39,	0.83 ± 0.16[f]	0.85 ± 0.14[f]
44 – 62, 67 – 92, and 101 – 134	1.50 ± 0.21[g]	1.49 ± 0.20[g]
residues 4 – 134	1.39 ± 0.25[f]	1.46 ± 0.26[f]
	2.13 ± 0.32[g]	2.18 ± 0.31[g]
Atomic rmsd [Å][h] *in comparison with*		
GlmS	1.02	–
mutB	1.83	–
metH	1.96	–

[a] In the first set of structure calculations, Φ and Ψ predictions of the TALOS program for residues Leu 23–Ala 30 were included; in the second set of structure calculations, the dihedral angle restraints of these residues were not included. [b] For comparison the numbers of restraints used for the structure calculation of ^{15}N-labeled MutS are given in parentheses.[24] [c] E_{L-J} = Lenard – Jones van der Waals energy calculated with the CHARMM force field.[62] [d] PROCHECK-NMR[30] statistics apply to the ordered regions of MutS, with residues 4 – 10, 23/26 – 39, 44 – 62, 67 – 92, 101 – 134. [e] Pairwise rmsd relative to mean coordinates from the superposition of backbone atoms (N, C^α, carbonyl C) and for all heavy atoms, respectively. [f] For backbone atoms. [g] For all heavy atoms. [h] Pairwise rmsd between backbone atoms (N, C^α, C=O) of MutS (mean coordinates) and mutB/metH, calculated from a superposition of backbone atoms participating in elements of secondary structure ($\alpha1-\alpha5$, $\beta1-\beta5$) as defined in the text.

consists of a five-stranded twisted parallel β sheet (β1: residues 4 – 10, β2: residues 33 – 39, β3: residues 56 – 62, β4: residues 86 – 92, β5: residues 114 – 118) with the strand order β2-β1-β3-β4-β5. The β sheet is encased by four (five) roughly parallel α helices (α1: residues 26(23) – 30, α2: residues 43 – 53, α3a/α3b: residues 66 – 71 and 74 – 81, α4: residues 101 – 111, α5: residues 123 – 135). Formally, the number of α-helices differs from those determined earlier with ^{15}N-labeled MutS[24] and ^{15}N-labeled GlmS[25] as helix α3 is subdivided into two α-helical parts (α3a/α3b) and a short (dynamic) α-helical segment can be defined for the "nascent" helix α1. The apparent length of the "nascent" helix depends on the inclusion of the TALOS analysis for the Φ and Ψ torsion angles (based on the available information on the sequence and on secondary ΔH^α, ΔC^α, ΔC^β, ΔCO, and ΔN^H chemical shifts) of this segment. If the TALOS analysis is omitted from the input restraints, then helix α1 is less well defined and comprises only the four residues Ser 26 – Ala 30 at the C-terminal end of this segment. Inclusion of the torsion angle predictions from a TALOS analysis indicated the "nascent" helix α1 to extend further and to include the residues Leu 23 to Ala 30.

In the resulting solution structure of MutS, the α helices are packed against both sides of the β sheet, which is built up by largely hydrophobic residues: Helices α2, α3a/α3b, and α4 are on one side of the β sheet, the "nascent" helix α1 and helix α5 are on its other side (Figures 6C, D). The higher number of NMR-derived restraints resulted in a more compact structure than that from ^{15}N-labeled MutS.[24] About 24% of the 137 MutS residues are involved in β sheets and about 39 – 42% take part in the formation of helical structure. About 10% of residues contribute to the well-structured ("lower") turns between the N-terminal ends of the β strands and the C-terminal ends of the α helices, as well as to the two residues that link α3a and α3b. A larger percentage (21 – 24%) belong to the ("upper") less well-structured loops linking the C-terminal ends of the β strands and the N-terminal ends of the α helices. These less well-defined regions also comprise residues Gly 12 – His 25, with the "B_{12}-binding motif".

The average root-mean-square deviation (rmsd) from the 15 best MutS structures is 0.83 ± 0.16 Å for the backbone atoms (N, C^α, carbonyl C), calculated with exclusion of the unstructured segments (Table 1), compared to 1.12 ± 0.4 Å obtained earlier for ^{15}N-labeled MutS.[24] Considering the well-structured parts of MutS, the analysis of the dihedral angles Φ and Ψ of the 15 best structures by using the program PROCHECK-NMR[30] shows that all of these angles fall within the allowed region of the Ramachandran plot.

Discussion

Solution structure of MutS

The B_{12}-binding subunits of Glm, MutS from *C. tetanomorphum*, and GlmS from *C. cochlearium*, were recently revealed by analysis of ^{15}N-labeled samples by heteronuclear NMR spectroscopy to be largely folded globular proteins in aqueous solution.[24, 25] The tertiary structures of both B_{12}-free apoproteins were found to represent a variant of the mononucleotide-binding "Rossmann fold" and to feature a β sheet with five parallel β strands, surrounded by four well-defined α helices and a dynamic "nascent" helix. In both apoproteins the sequence with residues 13 – 27 is the region that contains the conserved B_{12}-binding Asp-Xxx-His-Xxx-Xxx-Gly motif and which is conformationally disordered in the absence of coenzyme B_{12}. By comparison with the crystal structures of the holoenzymes metH and Mcm,[15, 17, 18] this dynamic segment, which includes the "nascent" helix, was suggested to play a crucial role in the recognition and binding of the base-off form of the corrinoid cofactor.[24] A recent NMR spectroscopic study of apomyoglobin characterized it as another largely folded globular protein, in which a single large segment becomes α-helically structured only upon binding of the

porphinoid cofactor heme, and uncovered a situation remarkably related to that in MutS.[31]

The precision of the final structures of $^{13}C,^{15}N$-labeled MutS is considerable (Table 1) and the structured parts of the final fifteen MutS backbone traces overlay with an rmsd of 0.83 ± 0.16 Å, compared to 1.12 ± 0.4 Å obtained earlier for uniformly ^{15}N-labeled MutS.[24] The refined solution structure of doubly $^{13}C,^{15}N$-labeled MutS showed some structural details, often reproducing remarkably those found in the crystal structures of the holoenzyme Glm.[8] With doubly labeled MutS, the four (five) α helices, which surround a central β sheet consisting of five parallel β strands, could be analyzed more precisely. The derived structure shows helix α3 to be split into two halves, α3a/α3b (Figure 5), similar to the situation in the crystal.[8] Furthermore, the section corresponding to the "nascent" helix is better defined and two major states are seen to equilibrate, one with a short, well-structured helical part with a single turn (residues Ser 26 – Ala 30), the other with the helix now extending over nearly twice as many residues (Ile 22 – Ala 30, Figures 6 C, D).

Several experimental lines of evidence indicated the disruption of helix α3 by H-bonding interactions of its backbone with amide hydrogen atoms of the side chain of Gln 43, which is located at the N terminus of helix α2. In the solution structure of MutS, the side chain of Gln 43 points toward helix α3 (Figure 5) where its amide group inserts, forming a gap in the α3 main chain in the region of residues Cys 71 – Leu 74. As a consequence, the H-bonding network of helix α3 is severely disturbed and the backbone amide functions of residues Gly 76 – Lys 72 are prevented from forming intrahelical H bonds with the carbonyl oxygen atoms of their $i-4$ predecessors. The irregular structure of this moiety of MutS is well defined by a high number of restraints.

The solution structure of MutS, as derived from the ^{15}N-labeled protein, showed the "nascent" helix α1 to undergo fast conformational interconversion between random coil and an α helix.[24] The term "nascent helix" was coined earlier for this type of dynamic structure.[32] In $^{13}C,^{15}N$-labeled MutS, all residues in the crucial segment (Ala 17 – Ala 30) gave rise to detectable amide signals. Only a very limited number of interresidue NOEs were observed, a pattern that is not consistent with a well-structured α helix. Medium- and long-range NOEs, which are typical for α helices,[29] were only detected for the residues Ser 26 to Ala 30, near the C-terminal end of the "nascent" helix. On this basis, a short helix α1 was calculated between residues Ser 26 and Ala 30, but the MutS residues between Gly 12 and His 25 were indicated to be less structured (Figures 4, 6 C). On the other hand, ΔH^{α} and ΔC^{α} secondary shifts of MutS residues Ile 22 – Asn 29 (Figure 3) are clearly consistent with an α-helical conformation of this section as well. Since, in the case of fast conformational exchange, chemical shift values represent an average across all the conformational states that the nuclei experience, residues Ile 22 – Asn 29 of MutS are indicated to exhibit significant (time-averaged) α-helical conformation. The TALOS database predicts nonambiguous Φ and Ψ backbone torsion angle restraints over all residues between Ile 22 and Ala 30, favoring an α-helical conformation of this segment. Indeed, inclusion of torsion angle restraints from TALOS in the structure calculations extends helix α1 from Ala 30 to residue Leu 23 (Figure 6 D). Accordingly, MutS is indicated to exist in an equilibrium between an $\alpha_5\beta_5$ fold, in which the larger part of helix α1 is present, and an $\alpha_4\beta_5$ fold, in which only the C-terminal part of this dynamic segment (Ser 26 – Ala 30) adopts an α-helical conformation. Interestingly, residues Leu 23 – Ala 30 (forming the "nascent" helix α1) have slightly smaller Chou-and-Fasman-type propensities (average: 0.98) for an α-helical conformation when compared with the residues in the MutS helices α2 – α5 (average: 1.12).[33] The structure models, shown in Figures 6 A/C and B/D, can thus be regarded as representing the two conformationally exchanging states of the MutS protein.

Specific interactions between secondary structure elements lining the nucleotide-binding cleft, that is, between specific side chains of residues participating in the "nascent helix" α1 and strand β2, as well as between helix α5 and strands β4/β5, appear to be relevant features of MutS. The (temporary) presence of an H bond between the carboxylate group of Asp 24 (of the "nascent" helix) and the side chain amide function of Asn 36 (in β2) is observed (Figure 7 A). This H bond is likely to exist only in the extended α-helical conformation of the "nascent" helix and thereby is suggested to stabilize the $\alpha_5\beta_5$ fold (Figure 6 D) of MutS with respect to its $\alpha_4\beta_5$ fold (Figure 6 C). In the $^1H,^{15}N$-HSQC spectrum, the side chain amide group of Asn 36 gives rise to cross peaks with the lowest intensity when compared to the corresponding cross peaks of the remaining Asn and Gln

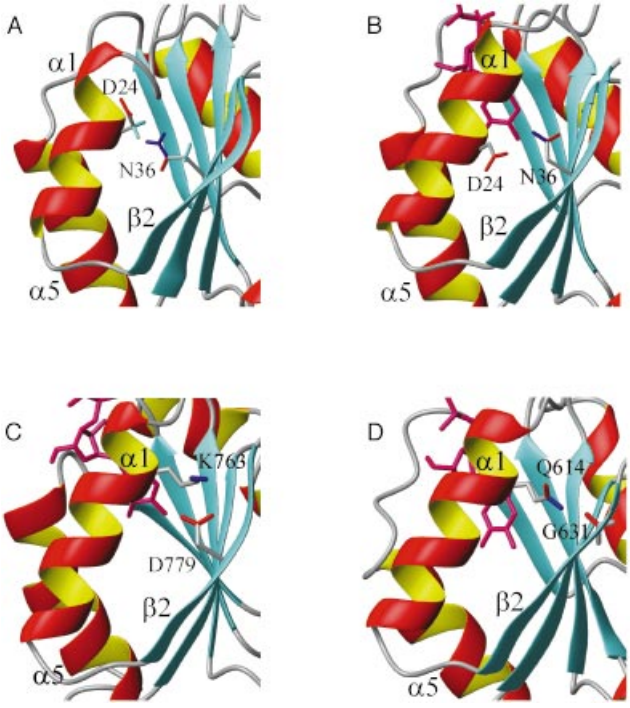


Figure 7. *Ribbon drawing close-ups of B_{12}-binding subunits showing side chain H-bonding interactions between residues in helix α1 and β strand β2. A: Representative MutS solution structure in the $\alpha_5\beta_5$ fold. B: Section of the B_{12}-binding subunit GlmS from the crystal structure of Glm from* C. cochlearium.[8] *C and D: Sections of the crystal structures of the B_{12}-binding domain of MetH from* E. coli[15] *and of Mcm from* P. shermanii.[17] *The side chains that are involved in H-bond formation are depicted as stick drawings and are labeled with their residue number. (The figures were generated with the program MOLMOL.*[63])

residues of MutS. This observation is consistent with the interpretation that these nuclei experience (additional) relaxation contributions from conformational exchange. Asp 24 and Asn 36 belong to the 84% of residues that are conserved between MutS and its functional analogue GlmS, and the presence of a hydrogen-bonding interaction between the side chains of Asp 24 and Asn 36 in the solution structure of GlmS is independently supported experimentally.[25] Hydrogen bonds between side chains of residues participating in helix α1 and β strand β2 are also observed in the available crystal structures of Glm and other B_{12}-dependent enzymes that bind their B_{12} cofactor in the "base-off/His-on" mode (see below).

Another H bond between secondary structure elements that is observed in the solution structure of MutS[24] (and similarly observed in GlmS from *C. cochlearium*[25]) concerns that between Tyr 117 (of strand β5) and Asp 129 (of helix α5) (Figure 8). In addition, in the structures of GlmS[25] and the Glm holoprotein[8] a second interstrand H bond is found between Tyr 89 of β4 and Asp 133 of helix α5. In contrast, a hydrophobic contact is seen in the solution structure of MutS between the corresponding "doubly exchanged" pair Phe 89 and Val 133. The available data point to the existence of specific interactions between the β sheet and helices α1 and α5, which are likely to be of particular relevance for the dynamic structuring of the nucleotide-binding cleft.

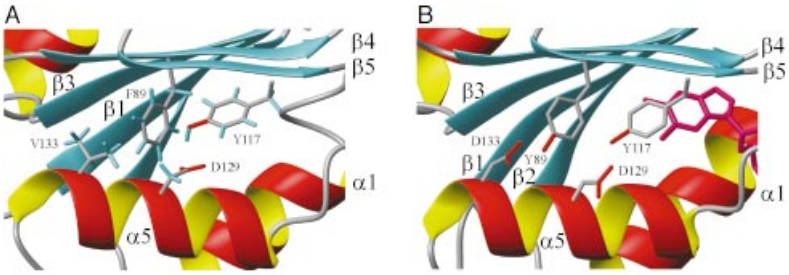


Figure 8. *Side chain hydrogen-bonding interactions at the interface of strands β4/β5 and helix α5 of B_{12}-binding subunits of Glm. A: Section of a representative solution structure of MutS in the $\alpha_5\beta_5$ fold. B: Section of the crystal structure of Glm from* C. cochlearium *with B_{12} bound.[8] A and B show the five-stranded β sheet (cyan) and helices α5 and α1 (red/yellow) of MutS and GlmS, respectively. In Figure A, the side chains of MutS residues Phe 89, Tyr 117, Asp 129, and Val 133 are depicted as stick drawings, in Figure B the side chains of GlmS residues Tyr 89, Tyr 117, Asp 129, and Asp 133, as well as the DMB base of the bound B_{12}. (The figures were generated by using the program MOLMOL.[63])*

Structural comparison of the solution structure of MutS with the crystal structures of Glm and of related B_{12}-dependent enzymes

Topological features: Ribbon models of the B_{12}-binding domains (subunits) from the crystal structures of the holoenzymes Mcm from *P. shermanii* and Glm from *C. cochlearium* as well as two solution structures of the B_{12}-binding subunit (MutS) from *C. tetanomorphum* are compared in Figure 6. The fold of the B_{12}-free subunit MutS in solution shares striking similarities with the crystal structures of the B_{12}-binding domains of Mcm and of Glm in the B_{12}-bound state. The apoprotein MutS and the B_{12}-binding domains of the holoenzymes Mcm and Glm represent variants of the "Rossmann fold", with a pattern of alternating β sheet and α helix structural elements. MutS shares a high degree of structural identity with its functional analogue GlmS (84% sequence identity) and the backbone heavy atoms of residues participating in elements of regular secondary structure (α1 – α5, β1 – β5) of MutS and GlmS superimpose with an rmsd of 1.02 Å (Table 1). The sequence identity of MutS with the B_{12}-binding domains of Mcm and metH only amounts to 31 and 19%, respectively, correlating with larger rmsd values (Table 1). The high degree of similarity of MutS and GlmS is partly due to some remarkable, common structural features (conformational irregularity in β strand β2; splitting of helix α3 into two halves (α3a/α3b) (Figures 5, 6C, D, F)). For Mcm (Figure 6E) and metH the conformational irregularity of β2 was also observed, but not a disruption of helix α3.

H-bonds crucial for tertiary structure: The H-bond between the side chains of Asp 24 and Asn 36, as observed in the solution structure of MutS (see Figure 7A), is clearly discernible in the crystal structure of Glm from *C. cochlearium*[8] (Figure 7B). In addition, the crystal structure features two more H-bonds from the side chain amide function of Asn 36 to residue Asp 20. Equivalent H-bonds are not deducible from the NMR spectroscopic data on MutS, where residue Asn 20 is indicated to be dynamic and to occupy random positions. The X-ray crystal structures of the B_{12}-binding domain (metH) of methylcobalamin-dependent MetH from *E. coli*[15] and of Mcm, the coenzyme B_{12}-dependent mutase from *P. shermanii*,[17] also feature a side chain hydrogen bond between α1 and β2, although the sequence identity of both proteins with MutS is much less significant (see above). The metH residue Asp 779 is located in β2 at a position comparable to that of Asn 36 in MutS (Figure 7C). The Asp 779 carboxylate group forms a hydrogen bond with the (protonated) amino group of Lys 763, located in helix α1. The position of Lys 763 in metH corresponds to the position of residue Asn 20 in the MutS/GlmS sequence. In the Mcm crystal structure (Figure 7D), a similar hydrogen bond between the Mcm residue Gln 614 (H-bond donor, in α1) and the backbone carbonyl oxygen atom of Gly 631 (H-bond acceptor, in β2) is again observed. Gly 631 is shifted two residues toward the C terminus of β2, when compared to the position of Asn 36 in MutS/GlmS and Asp 779 in metH. A corresponding glycine (Gly 631) is remarkably well conserved in all four sequences of B_{12}-binding proteins (Gly 38 in MutS/GlmS, Gly 781 in metH) and the observed H-bonds between α1 and β2 (Asp 24 – Asn 36 in Glm, Lys 763 – Asp 779 in metH, Gln 614 – Gly 631 in Mcm) are likely to play a role in the stabilization of helix α1 and in populating the more extended helical conformation of the "nascent" helix of MutS.

Likewise, specific pairs of residues of helix α5 and strands β4 and β5 of the β sheet were observed to interact in the solution structures of MutS and GlmS, as was found in the crystal structure of the Glm holoenzyme (Figure 8). Interestingly, half of the residues that are nonconserved between GlmS and MutS

form such pairs with complementary side chains that mutually interact at the interface of the strands β3 and β4 and helix α5.[25] A set of specific stabilizing interactions between the β sheet and the helices α1 and α5 are thus indicated to help preorganize the nucleotide-binding cleft in the B_{12}-binding subunits MutS and GlmS.

Tertiary structure: Detailed information about the degree of similarity of the structures of MutS in solution and of GlmS (with B_{12} bound) in the crystal on a per residue basis is obtained by inspection of the pairwise rmsd values (Figure 9). With the exception of helix α3b, (average rmsd values of 1.3 Å), the remaining elements of regular secondary structure match better with an average pairwise rmsd of 0.6 – 1.1 Å. The important structural differences concern the "nascent" helix α1 in MutS and GlmS. While the calculated relative position of this section of MutS is not yet well defined at its N-terminal part (due to lack of long-range side chain NOEs of this dynamic section), the differing apparent length of the "nascent" helix is of interest.

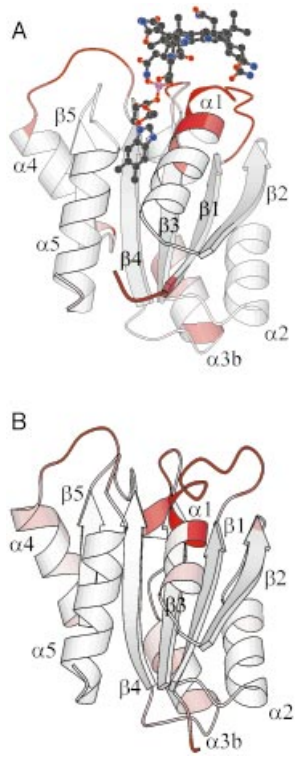


Figure 9. *Comparison of the B_{12}-free MutS solution structure(s) and the GlmS subunit (with B_{12} bound) from the Glm crystal structure.[8] A: Pairwise rmsd values between the Glm X-ray crystal structure and the MutS solution structure (mean coordinates for $\alpha_4\beta_5$ fold of MutS) were calculated from a superposition of backbone N, C^{α}, and carbonyl C atoms, based on defined elements of secondary structure in GlmS (α1 – α5, β1 – β5). Color code: white = rmsd ≤ 1.5 Å; intermediate intensities of red = 1.5 Å ≤ rmsd ≤ 2.5 Å; dark red = rmsd ≥ 2.5 Å. B: Pairwise rmsd values between two representative solution structures of the $\alpha_4\beta_5$ fold and the $\alpha_5\beta_5$ fold of MutS, calculated from a superposition of backbone N, C^{α}, and carbonyl C atoms, based on defined elements of secondary structure in MutS (α1 – α5, β1 – β5). Color code: white = rmsd ≤ 0.5 Å; intermediate intensities of red = 0.5 Å ≤ rmsd ≤ 1.5 Å; dark red = rmsd ≥ 1.5 Å. (The figures were generated using the program MOLSCRIPT.[64])*

The MutS structures shown in Figures 6 A and B indicate the "nascent" helix α1 to be composed of about eight residues (Leu 23 – Ala 30). The crystal structure of Glm (Figure 6 F) shows helix α1 to be still longer (Val 18 – Ala 30) and to comprise five more residues at its N-terminal end. The Glm crystal structure shows helix α1 to form part of the nucleotide-binding cleft and to furnish the nonpolar residues (Leu 23 and Phe 27 in GlmS) for contact with the B_{12} nucleotide.

The loops: The MutS segments Ile 11 – Ile 22, Asn 93 – His 100 and Pro 119 – Pro 122 form the "upper" loops that link the N termini of the α helices with the C termini of the β strands. The loop between Ile 11 – Ile 22 harbors the invariant B_{12}-binding motif Asp-Xxx-His-Xxx-Xxx-Gly. In the crystal structure of Glm from *C. cochlearium* reconstituted with methylcobalamin in the "base-off/His-on" form, His 16 of this conserved motif coordinates to the cobalt center and is H-bridged to Asp 14, while the B_{12} nucleotide tail of the B_{12} cofactor is buried in the binding cleft of the protein.[8] The "upper" loops thus assist in the binding of the B_{12} coenzyme and strongly interact with the protein-bound corrinoid cofactor: The amide H atoms of residues Cys 15, Ala 17, Val 18, Gly 19, Gly 97, Gly 120, and Thr 121, carbonyl oxygen atoms of Ser 13 and Val 95, and the side chain oxygen atom of Asn 93 form direct or water-relayed hydrogen bonds to the propionamide side chains, the phosphate group, and/or the ribose unit of the B_{12} cofactor in the Glm crystal.[8] In the MutS solution structure, the loops Ile 11 – Ile 22 and Asn 93 – His 100 appear disordered as also observed for most of the "nascent" helix α1 (Figure 4). Previous ^{15}N-relaxation data[24] showed this disorder to result from the intrinsic flexibility of these parts of MutS and to occur over a wide time scale: Residues between Ser 13 and Gly 19 are indicated to be involved in fast motions (picoseconds to nanoseconds), whereas residues between Ile 22 – Thr 28 and Gly 92 – Gly 97 appear to be mobile with motions occurring on the 100-μs time scale.

The nucleotide-binding cleft: An overlay of the ribbon models of the MutS solution structure with the Glm crystal structure with B_{12} bound (Figure 9 A) shows the apoprotein MutS to be largely preorganized (even to the extent of several structural "irregularities") for binding of the B_{12} coenzyme in a "base-off/His-on" mode. The empty nucleotide-binding cleft in the B_{12}-free MutS solution structure is preformed for binding the B_{12} cofactor with the exception of the "nascent" helix (Figure 9 B). The B_{12} cofactor, as observed in the Glm crystal structure,[8] can be fitted into the preformed cleft without severe steric clashes (Figure 10). A majority of the MutS side chains, which are expected to form the hydrophobic binding cleft, already adopt orientations resembling those observed in the Glm holoenzyme crystal structure. The lipophilic side chains of Leu 23 and Leu 63 are conformationally undefined in MutS in the absence of the B_{12} cofactor, but are positioned to make hydrophobic contacts with the nucleotide moiety of the latter (Figure 10). The hydroxymethylene side chain of Ser 61 (which is hydrogen-bonded to N3 of the bound DMB base in the holoenzyme) extends from β3 at the center of the β sheet and would already be in place to help to fix the

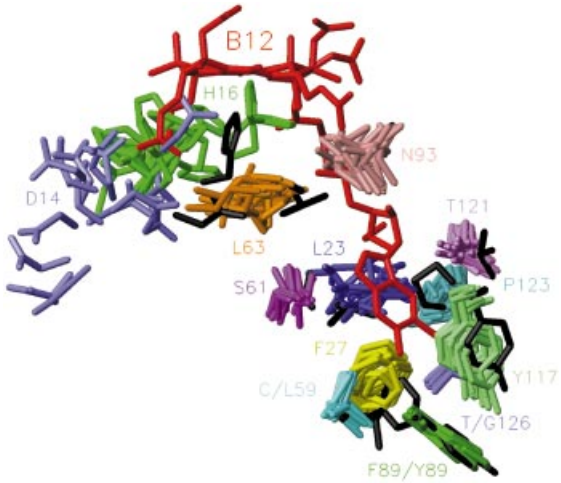


Figure 10. *Analysis of B_{12} binding to the B_{12}-binding subunit of glutamate mutase. GlmS side chains interacting with the B_{12} cofactor according to the Glm crystal structure (black) with bound B_{12} (red) are superimposed onto the corresponding side chains (shown in various other colors) from the 15 best NMR spectroscopically derived solution structures of the $\alpha_5\beta_5$ fold of MutS. (The figure was generated using the program MOLMOL.[63])*

position of the base of the α-configurated pseudonucleotide of coenzyme B_{12}.

MutS selectively binds the detached nucleotide moiety of B_{12} in the nucleotide-binding cleft in an $\alpha_5\beta_5$ topology,[26] which is similar to that of the holoenzyme crystal structures.[8] The bound nucleotide, in turn, stabilizes the $\alpha_5\beta_5$ structure of the protein. The residues of the conserved B_{12}-binding motif are dynamic in the apoprotein MutS and remain so in MutS carrying the B_{12} nucleotide moiety.[26] The $\alpha_5\beta_5$ state of MutS thus is dynamically preorganized to trap the nucleotide moiety of the base-off form of the coenzyme. Both the B_{12}-binding protein MutS and the bound B_{12} cofactor are indicated to adopt their structures mutually. Work is currently in progress in our laboratories to delineate a dynamic mechanism for the resulting "induced fit" (see for example ref. [34]) of these two components, based on the well-defined solution structure of MutS, as available from the present study.

Experimental Section

Preparation of uniformly ^{13}C,^{15}N-labeled MutS protein:

Uniformly ^{13}C,^{15}N-doubly-labeled MutS protein was prepared by growing *E. coli* BL21 (DE3) harboring the plasmid pmutSX[12] on M9 minimal medium supplemented with 2 g L^{-1} ^{13}C-labeled glucose as the carbon source and 0.66 g L^{-1} ^{15}N-labeled ammonium chloride as the nitrogen source. Ampicillin (100 mg L^{-1}) was included to maintain selection for the plasmid. Cultures were grown at 37 °C and expression of *mutS* induced by the addition of 200 mg L^{-1} isopropyl-β-thiogalactoside (IPTG) when cells reached an optical density (measured at 600 nm; OD_{600}) of 1.0. The cells were allowed to grow to stationary phase and harvested by centrifugation. MutS was purified as described previously.[12]

NMR spectroscopic measurements and data processing: All NMR experiments were performed on a Varian UNITY Plus 500 MHz spectrometer equipped with a pulsed-field gradient unit triple-resonance probe with actively shielded *z* gradients. The sample was kept in 250-µL Shigemi tubes and contained 1.3 mM MutS, 11 mM potassium phosphate, pH 6.0, 0.5 mM ethylenediamine tetraacetate (EDTA), 5 mM dithiothreitol, 1.5 mM NaN_3 in H_2O/D_2O (9:1). All spectra were recorded at 26 °C. Experiments involving amide proton detection used pulsed-field gradients for coherence transfer pathway selection[35–37] making use of an enhanced sensitivity approach[38, 39] with minimal H_2O saturation and dephasing.[40–42] The following experiments were used in the present study for spin system identification and sequential assignment: two-dimensional (2D) ^{15}N heteronuclear single-quantum correlation (HSQC)[35, 43] 2D (HB)CB(CGCD)HD,[44] 2D (HB)CB(CGCDCE)HE,[44] three-dimensional (3D) HNCO,[42] 3D HNCA,[45] 3D HNCACB,[46] 3D CBCA(CO)NH,[47] 3D HCCH-TOCSY,[48] and 3D ^{13}C NOESY-HSQC (NOESY = nuclear Overhauser effect spectroscopy).[49] The HCCH-TOCSY (TOCSY = total correlation spectroscopy) was acquired using decoupling in the presence of scalar interactions (DIPSI-2)[50] for carbon isotropic mixing. All triple-resonance spectra were performed including a water flip-back pulse to minimize the effects of radiation damping and solvent exchange.[40, 41] Typical carrier positions employed in the double- and triple-resonance experiments were $\delta = 116$ for ^{15}N, $\delta = 178$ for ^{13}CO, $\delta = 58$ for $^{13}C^{\alpha}$, $\delta = 43$ for $^{13}C^{\alpha}/^{13}C^{\beta}$ (i.e. aliphatic carbons), and $\delta = 4.68$ for ^{1}H. Experiments in which protons directly bound to nitrogen atoms were detected during acquisition employed wide-band uniform rate and smooth truncation (WURST) ^{15}N decoupling,[51] whereas experiments which detected aliphatic protons during acquisition made use of GARP for ^{13}C decoupling.[52] Where necessary (e.g. in HNCACB experiments), carbonyl decoupling was achieved using a SEDUCE-1 ^{13}C decoupling sequence centered at $\delta = 178$.[53] Quadrature detection in all of the indirectly detected dimensions was achieved by States-TPPI.[54] Spectra were processed using NMRPipe software[55] and visualized and assigned using ANSIG software.[56] The number of data points was doubled by either mirror image linear prediction (for constant time evolution) or forward–backward linear prediction (nonconstant time evolution) in the indirect dimensions prior to zero-filling and Fourier transformation (FT).

Input restraints: The interproton distances were estimated from nuclear Overhauser enhancement (NOE) intensities observed in 3D ^{1}H,^{13}C NOESY-HSQC spectrum. Further NOE restraints were obtained from 2D SS NOESY (SS = symmetrically shifted shaped pulses)[56] and 3D ^{1}H,^{15}N NOESY-HSQC spectra, which have been recorded and used for solving the solution structure of ^{15}N-labeled MutS.[24] Prior to a renewed integration of these two spectra, the NOE assignment therein was completed by using the full signal assignment provided by the doubly labeled MutS sample. Cross peaks were integrated with ANSIG software.[57] NOE intensities were calibrated to distances on the basis of short interproton distances observed in elements of regular secondary structure or by using covalently fixed distances. They were classified as strong (1.8–2.7 Å), medium (1.8–3.3 Å), weak (1.8–5 Å), and very weak (1.8–6 Å). For nonstereospecifically assigned or equivalent protons, NOE distances were assigned as a $(\Sigma r^{-6})^{-1/6}$ sum.[58]

Backbone dihedral angles Φ and Ψ were obtained by employing TALOS software,[28] that is, a database system for empirical prediction of Φ and Ψ angles using a combination of five kinds ($^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, ^{13}CO, ^{15}N) of chemical shift assignments for a given protein sequence. According to the empirical "rules of thumb" implemented in the TALOS software (TALOS = torsion angle likelihood obtained from shift and sequence similarity),[28] for 78 MutS residues the Φ and Ψ predictions were classified to be good (unambiguous) and were consequently used as dihedral angle restraints in subsequent

structure calculations. In the later stages of structure calculation, Φ and Ψ predictions, which were classified by the TALOS program to be ambiguous (but not necessarily wrong), were inserted into the constraints input file for additional 13 MutS residues (see "Structure calculations" below). The upper and lower bounds for the Φ and Ψ target values were set to 13° (Ψ) and 14° (Φ); these values are reported to represent on average the uncertainty of good TALOS predictions.[28] Exceptions to this were some Φ and Ψ angles for which TALOS reported higher standard deviations than 13° and 14°, respectively.

Hydrogen bond restraints between amide protons and their carbonyl acceptors in elements of regular secondary structure were defined and used in structure calculations as in the case of ^{15}N-labeled MutS.[24] The only major exception to this were residues Lys 72 – Glu 76, located in disrupted helix α3, for which no such restraints were applied (see "Results"). No hydrogen bond restraints were defined for the conformationally flexible "nascent" helix α1,[24] which consists of the amino acids Val 18 – Ala 30. All hydrogen bonds in elements of regular secondary structure were explicitly defined as distance restraints between backbone amide protons and carbonyl oxygen atoms, corresponding to the limits of 1.8 – 2.5 Å. Additional hydrogen bonds were detected between the hydroxy group proton of Tyr 117 and the carboxyl group of Asp 129 as well as between one of the side chain amide protons of Asn 36 and the carboxylate group of Asp 24. The presence of these H-bonds can be deduced from the spectral observation of the corresponding exchange-labile side chain protons, indicating protection from exchange with bulk water protons, and from the extreme low-field shift of these protons (Tyr 117: $\delta = 11.1$, Asn 36: $\delta = 8.65$). In the case of Tyr 117, the identification of the carboxylate function of Asp 129 as the most likely H-bond acceptor was supported by the observation of NOEs between the hydroxy group proton of Tyr 117 and the side chain protons H$^{\beta}$ of Asp 129. Likewise, an NOE is observed between the lowfield-shifted side chain amide hydrogen atom of Asn 36 and the H$^{\beta}$ protons of Asp 24. Both side chain – side chain H-bonds (Asp 24/Asn 36, Tyr 117/Asp 129) were used as distance restraints, defined with the limits of 1.8 – 2.8 Å.

Structure calculations: Three-dimensional structures were generated according to standard distance geometry (DG)[59]/restrained simulated annealing (rSA)[60]/simulated annealing (SA) refinement/energy minimization (EM) protocols using X-PLOR software (version 3.8.1)[61] running on a cluster of Silicon Graphics workstations. First, an initial ensemble of 163 structures was calculated from a covalent template structure with randomized backbone dihedral angles Φ and Ψ and extended side chains. In this first stage, a total of 1483 NOE-derived distance restraints, Φ and Ψ dihedral angle restraints for 78 MutS residues as well as 55 hydrogen bond restraints were applied. 100 structures with a minimum of restraint violations and minimal energy were selected for further iterations of rSA/SA refinement/EM. In the course of this process, more NOE restraints could be identified and introduced into the calculations. In addition, after every iteration of rSA/SA refinement/EM the backbone dihedral angles of the resulting structures were compared with their corresponding empirical TALOS prediction values, in particular the backbone dihedral angles of those residues whose predictions have been classified as ambiguous (but not necessarily wrong) in an initial TALOS output. If these Φ and Ψ prediction values were in the order of the corresponding values in the calculated structures ($\pm 25°$), they were added to the input restraints set and used in subsequent iterations. Finally, the resulting 100 structures were minimized until convergence by use of the CHARMM force field (CHARMM = chemistry at Harvard macromolecular mechanics)[62] employing about 1800 restraints (from 1553 NOEs, 55 H-bonds, and 184 Φ and Ψ backbone dihedral angles).

A second run of structure calculations was carried out in a single iteration generating further 100 MutS structures from a covalent template structure with randomized backbone dihedral angles Φ and Ψ and extended side chains according to DG/rSA/SA refinement/EM protocols. In these structure calculations the constraint input file differed from that of the first run by omitting the Φ and Ψ torsion angle restraints for the residues of the "nascent" helix α1.[24] All structures were subjected to a final refinement using the CHARMM force field. From each family of MutS structures (resulting from the first and second run of structure calculations), a final set of 15 structures with a minimum of restraint violations and minimal energy were selected for structural analysis.

Accession number: The coordinates of the solution structure of MutS have been deposited in the Brookhaven Protein Data Bank (PDB) under the accession code 1FMF.

We thank Georg Kontaxis and Christian Eichmüller (Institute of Organic Chemistry, University of Innsbruck) for their expertise and advice. We are grateful to the Austrian National Science Foundation (FWF project no. P-13595), the European Commission (TMR project no. FMRX.CT96.0018), and the National Institutes of Health (grant GM 59227 to E.N.G.M.) for financial support and would like to acknowledge the funding of the major equipment by the Austrian Ministry of Science.

[1] H. A. Barker, V. Rooze, F. Suzuki, A. A. Iodice, *J. Biol. Chem.* **1964**, *239*, 3260 – 3266.
[2] U. Leutbecher, R. Böcher, D. Linder, W. Buckel, *Eur. J. Biochem.* **1992**, *205*, 759 – 765.
[3] W. Buckel, H. A. Barker, *J. Bacteriol.* **1974**, *117*, 1248 – 1260.
[4] W. Buckel, *Arch. Microbiol.* **1980**, *127*, 167 – 169.
[5] R. L. Switzer in *B_{12}, Vol. 2* (Ed.: D. Dolphin), Wiley, New York, **1982**, pp. 289 – 305.
[6] O. Zelder, B. Beatrix, U. Leutbecher, W. Buckel, *Eur. J. Biochem.* **1994**, *226*, 577 – 585.
[7] H.-P. Chen, E. N. G. Marsh, *Biochemistry* **1997**, *36*, 7884 – 7889.
[8] R. Reitzer, K. Gruber, G. Jogl, U. G. Wagner, H. Bothe, W. Buckel, C. Kratky, *Structure* **1999**, *7*, 891 – 902.
[9] E. N. G. Marsh, D. E. Holloway, *FEBS Lett.* **1992**, *310*, 167 – 170.
[10] D. E. Holloway, E. N. G. Marsh, *FEBS Lett.* **1994**, *317*, 44 – 48.
[11] O. Zelder, B. Beatrix, W. Buckel, *FEMS Microbiol. Lett.* **1994**, *118*, 15 – 22.
[12] D. E. Holloway, E. N. G. Marsh, *J. Biol. Chem.* **1994**, *269*, 20425 – 20430.
[13] E. N. G. Marsh, D. E. Holloway, H.-P. Chen in *Vitamin B_{12} and B_{12} Proteins* (Eds.: B. Kräutler, B. T. Golding, D. Arigoni), Wiley-VCH, Weinheim, **1998**, pp. 253 – 264.
[14] B. Beatrix, O. Zelder, D. Linder, W. Buckel, *Eur. J. Biochem.* **1994**, *221*, 101 – 109.
[15] C. L. Drennan, S. Huang, J. T. Drummond, R. G. Matthews, M. L. Ludwig, *Science* **1994**, *266*, 1669 – 1674.
[16] M. L. Ludwig, R. G. Matthews, *Annu. Rev. Biochem.* **1997**, *66*, 269 – 313.
[17] F. Mancia, N. H. Keep, A. Nakagawa, P. F. Leadlay, S. McSweeney, B. Rasmussen, R. Bösecke, O. Diat, P. R. Evans, *Structure* **1996**, *4*, 339 – 350.
[18] P. R. Evans, F. Mancia in *Vitamin B_{12} and B_{12} Proteins* (Eds.: B. Kräutler, B. T. Golding, D. Arigoni), Wiley-VCH, Weinheim, **1998**, pp. 217 – 226.
[19] *Introduction to Protein Structure* (Eds.: C. Branden, J. Tooze), Garland Publishing, New York, **1991**.
[20] H. A. Barker, H. Weissbach, R. D. Smith, *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 1093 – 1097.
[21] B. Hoffmann, M. Oberhuber, E. Stupperich, H. Bothe, W. Buckel, R. Konrat, B. Kräutler, *J. Bacteriol.* **2000**, *182*, 4773 – 4782.
[22] J. N. Ladd, H. P. C. Hogenkamp, H. A. Barker, *J. Biol. Chem.* **1961**, *236*, 2114 – 2118.

[23] B. Kräutler, C. Kratky, *Angew. Chem.* **1996**, *108*, 179 – 183; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 168 – 172.
[24] M. Tollinger, R. Konrat, B. H. Hilbert, E. N. G. Marsh, B. Kräutler, *Structure* **1998**, *6*, 1021 – 1033.
[25] B. Hoffmann, R. Konrat, H. Bothe, W. Buckel, B. Kräutler, *Eur. J. Biochem.* **1999**, *263*, 178 – 188.
[26] M. Tollinger, C. Eichmüller, R. Konrat, M. Huhta, E. N. G. Marsh, B. Kräutler, *J. Mol. Biol.* **2001**, *309*, 777 – 791.
[27] D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **1991**, *222*, 311 – 333.
[28] G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **1999**, *13*, 289 – 302.
[29] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, **1986**.
[30] R. A. Laskowski, M. W. McArthur, D. S. Moss, J. M.Thornton, *J. Appl. Crystallogr.* **1991**, *24*, 946 – 950.
[31] D. Eliezer, J. Yao, H. J. Dyson, P. E. Wright, *Nat. Struct. Biol.* **1998**, *5*, 148 – 155.
[32] H. J. Dyson, M. Rance, R. A. Houghten, P. E. Wright, R. A. Lerner, *J. Mol. Biol.* **1988**, *201*, 201 – 217.
[33] M. B. Swindells, M. W. McArthur, J. M. Thornton, *Nat. Struct. Biol.* **1995**, *2*, 569 – 603.
[34] A. Fersht, *Enzyme Structure and Mechanism*, Freeman, New York, **1985**.
[35] L. E. Kay, P. Keifer, T. Saarinen, *J. Am. Chem. Soc.* **1992**, *114*, 10663 – 10665.
[36] J. Schleucher, M. Sattler, C. Griesinger, *Angew. Chem.* **1993**, *105*, 1518 – 1521; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1489 – 1492.
[37] D. R. Muhandiram, L. E. Kay, *J. Magn. Reson. Ser. B* **1994**, *103*, 203 – 216.
[38] J. Cavanagh, A. G. Palmer, P. E. Wright, M. Rance, *J. Magn. Reson.* **1991**, *91*, 429 – 436.
[39] A G. Palmer, J. Cavanagh, P. E. Wright, M. Rance, *J. Magn. Reson.* **1991**, *93*, 151 – 170.
[40] S. Grzesiek, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 12593 – 12594.
[41] J. Stonehouse, G. L. Shaw, J. Keeler, E. D. Laue, *J. Magn. Reson. Ser. A* **1994**, *107*, 178 – 184.
[42] L. E. Kay, G.-Y. Xu, T. Yamazaki, *J. Magn. Reson. Ser. A* **1994**, *109*, 129 – 133.
[43] G. Bodenhausen, D. H. Ruben, *Chem. Phys. Lett.* **1980**, *69*, 185 – 189.
[44] T. Yamazaki, J. D. Forman-Kay, L. E. Kay, *J. Am. Chem. Soc.* **1993**, *115*, 11054 – 11055.
[45] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *96*, 432 – 440.
[46] M. Wittekind, L. Mueller, *J. Magn. Reson. Ser. B* **1993**, *101*, 201 – 205.
[47] S. Grzesiek, A. Bax, *J. Am. Chem. Soc.* **1992**, *114*, 6291 – 6293.
[48] L. E. Kay, G.-Y. Xu, A. U. Singer, D. R. Muhandiram, J. D. Forman-Kay, *J. Magn. Reson. Ser. B* **1993**, *101*, 333 – 337.
[49] D. R. Muhandiram, N. A. Farrow, G.-Y. Xu, S. H. Smallcombe, L. Kay, *J. Magn. Reson. Ser. B* **1993**, *102*, 317 – 321.
[50] S. P. Rucker, A. J. Shaka, *Mol. Phys.* **1989**, *68*, 509 – 517.
[51] E. Kupce, R. Freeman, *J. Magn. Reson. Ser. A* **1995**, *115*, 273 – 276.
[52] A. J. Shaka, P. B. Barker, R. Freeman, *J. Magn. Reson.* **1985**, *64*, 547 – 552.
[53] M. A. McCoy, L. Mueller, *J. Am. Chem. Soc.* **1992**, *114*, 2108 – 2112.
[54] D. Marion, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson.* **1989**, *85*, 393 – 399.
[55] F. Delaglio, *NMRPipe System of Software*, National Institutes of Health, Bethesda, MD, **1993**.
[56] S. Smallcombe, *J. Am. Chem. Soc.* **1993**, *115*, 4776 – 4785.
[57] P. J. Kraulis, *J. Magn. Reson.* **1989**, *24*, 627 – 633.
[58] M. Nilges, *Proteins* **1993**, *17*, 295 – 309.
[59] J. Kuszewski, M. Nilges, A. T. Brünger, *J. Biomol. NMR* **1992**, *2*, 33 – 56.
[60] M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* **1988**, *229*, 317 – 324.
[61] A. T. Brünger, *X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven, CT, USA, **1992**.
[62] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 635 – 637.
[63] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51 – 55.
[64] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946 – 950.

Received: January 25, 2001 [F 186]