doi:10.1006/jmbi.2001.5095 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **313**, 1103–1116

JMB

Crystal Structure of a Protein Repair Methyltransferase from *Pyrococcus furiosus* with its L-Isoaspartyl Peptide Substrate

Scott C. Griffith[1], Michael R. Sawaya[2], Daniel R. Boutz[1], Nitika Thapar[1] Jonathan E. Katz[1], Steven Clarke[1] and Todd O. Yeates[1,2]*

[1]*Department of Chemistry and Biochemistry and Molecular Biology Institute*

[2]*UCLA-DOE Laboratory of Structural Biology and Molecular Medicine, University of California Los Angeles Box 951569, Los Angeles CA 90095-1569, USA*

**Corresponding author*

Protein L-isoaspartyl (D-aspartyl) methyltransferases (EC 2.1.1.77) are found in almost all organisms. These enzymes catalyze the *S*-adenosylmethionine (AdoMet)-dependent methylation of isomerized and racemized aspartyl residues in age-damaged proteins as part of an essential protein repair process. Here, we report crystal structures of the repair methyltransferase at resolutions up to 1.2 Å from the hyperthermophilic archaeon *Pyrococcus furiosus*. Refined structures include binary complexes with the active cofactor AdoMet, its reaction product *S*-adenosylhomocysteine (AdoHcy), and adenosine. The enzyme places the methyl-donating cofactor in a deep, electrostatically negative pocket that is shielded from solvent. Across the multiple crystal structures visualized, the presence or absence of the methyl group on the cofactor correlates with a significant conformational change in the enzyme in a loop bordering the active site, suggesting a role for motion in catalysis or cofactor exchange. We also report the structure of a ternary complex of the enzyme with adenosine and the methyl-accepting polypeptide substrate VYP(L-isoAsp)HA at 2.1 Å. The substrate binds in a narrow active site cleft with three of its residues in an extended conformation, suggesting that damaged proteins may be locally denatured during the repair process in cells. Manual and computer-based docking studies on different isomers help explain how the enzyme uses steric effects to make the critical distinction between normal L-aspartyl and age-damaged L-isoaspartyl and D-aspartyl residues.

Keywords: protein repair; methyltransferase; *S*-adenosylmethionine; D-amino acids; protein isomerization

Introduction

As proteins age, their aspartyl and asparaginyl residues deamidate, isomerize, and racemize to produce D and L-isoaspartyl and D and L-aspartyl residues.[1–3] The major product is the L-isoaspartyl residue, whose extra methylene group in the polypeptide backbone causes a kink in the affected protein.[4,5] This damage can result in the loss of activity of a protein or in altered sensitivity to proteolysis.[6–9] However, L-isoaspartyl residues are recognized by a widely distributed enzyme, the protein L-isoaspartyl (D-aspartyl) methyltransferase (EC 2.1.1.77), which transfers a methyl group from *S*-adenosylmethionine (AdoMet) onto the carboxyl group of the unusual side-chain.[10] The methyl ester thus formed can spontaneously cyclize with the protein backbone to form the L-succinimide derivative. Then, depending on which bond is broken, this cyclic intermediate spontaneously hydrolyzes to give either the natural L-aspartyl product or the damaged L-isoaspartyl derivative again. Repeated rounds of the cycle can drive the conversion of L-isoaspartyl residues fully to L-aspartyl residues, thereby constituting a protein repair pathway for age-damaged proteins (Figure 1).[11–13] The protein L-isoaspartyl methyltransferase is able to drive this pathway in the direction of net repair by acting

Abbreviations used: AdoMet, *S*-adenosylmethionine; AdoHcy, *S*-adenosylhomocysteine; rms, root-mean-square.

E-mail address of the corresponding author: yeates@mbi.ucla.edu

0022-2836/01/051103–14 $35.00/0

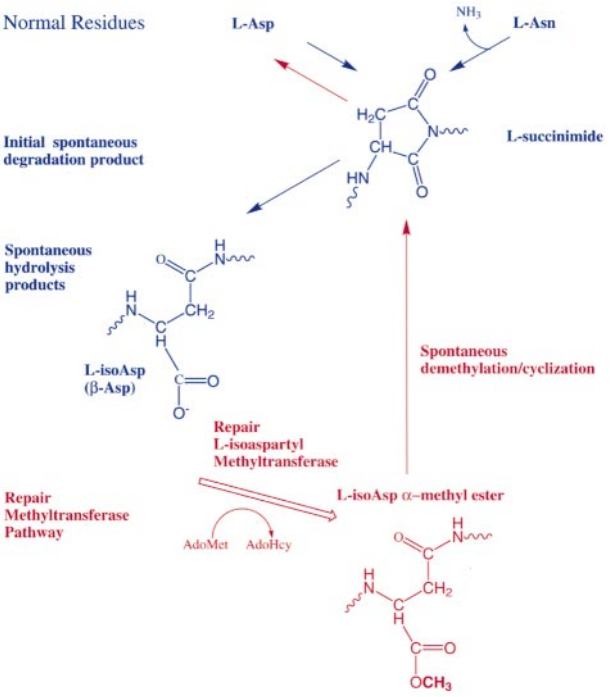


Figure 1. Spontaneous degradative pathways that generate abnormal L-isoaspartyl residues (blue) and the methyltransferase-dependent pathway involved in their conversion to natural L-aspartyl residues in a net repair reaction (red). The initial succinimide intermediate formed spontaneously from L-aspartyl and L-asparaginyl residues is non-enzymatically hydrolyzed to give both L-aspartyl residues and damaged L-isoaspartyl (β-aspartyl) residues that give a kinked polypeptide backbone. In a step that initiates the repair process, the L-isoaspartyl side-chain is recognized and methylated by the protein repair L-isoaspartyl (D-aspartyl) *O*-methyltransferase which catalyzes the transfer of a methyl group from the AdoMet cofactor, generating the α-methyl ester and AdoHcy. The methyl ester is itself unstable and is rapidly converted back to the succinimide, which can then spontaneously form either the normal L-aspartyl residue or regenerate the L-isoaspartyl residue. In the first case, the configurational damage is reversed. In the second case, repeated cycles of methylation and demethylation can result in the eventual conversion of the remaining L-isoaspartyl residues to L-aspartyl residues.[11–13] A similar repair pathway (involving racemization of the succinimide intermediate) can be drawn for the conversion of D-aspartyl residues to L-aspartyl residues.[15] Although the loss of the amide nitrogen atom from asparagine residues cannot be rectified, the same repair pathway does remove the kinks in the protein backbone that arise from the isomerization of both aspartyl and asparaginyl residues.

stereospecifically and by coupling its reaction to the energetically favorable demethylation of the AdoMet cofactor. Because the succinimide intermediate is also racemization-prone,[14] the conversion of unnatural D-aspartyl residues to L-aspartyl residues is also possible.[15]

The importance of this repair enzyme is made clear by its wide distribution, its high degree of amino acid sequence conservation and its required presence for a full life-span. L-Isoaspartyl methyltransferases have been found in nearly all organisms except Gram-positive bacteria, some yeasts, and some non-seed plants.[10,16–18] The genes encoding these enzymes are very stable, with an evolutionary mutation rate nearly as low as that of the cytochrome *c* protein family.[16] The survival of bacteria, worms, and mammals deficient in this enzyme is significantly reduced,[19–22] most notably in mice, where death occurs after only 5% of their normal life span.[23,24]

The high level of sequence conservation of this enzyme likely reflects the requirement that it must accurately discriminate between its major substrate, the age-damaged L-isoaspartyl residue, and the natural L-aspartyl residue. These two amino acids have few distinguishing features when examined within the context of a polypeptide chain. If the enzyme did methylate both natural L-aspartyl and damaged L-isoaspartyl residues, it would be unable to accomplish a net repair and would only hasten the formation of unnatural isomers. It has thus been of interest to elucidate the mechanism for this discrimination. It has also been of interest to understand how the enzyme can recognize one of the racemized isomers, the D-aspartyl residue, although this occurs at reduced affinity.[15,25]

Recently, it has been possible to begin to understand how the function of this enzyme is related to its structure. By amino acid sequence analysis, it is clear that the enzyme belongs to a large family of AdoMet-dependent small molecule, RNA, DNA, and protein methyltransferases.[26,27] In the last several years, the three-dimensional structures obtained for a number of AdoMet-dependent methyltransferase enzymes have shown that they share a common fold and belong to a structural superfamily.[28,29] Very recently, the first view of an L-isoaspartyl methyltransferase was provided from the eubacterial thermophile *Thermatoga maritima* as a complex with the demethylated cofactor *S*-adenosylhomocysteine (AdoHcy).[30] Here, we report the three-dimensional structure of an L-isoaspartyl methyltransferase from the archaeal thermophile, *Pyrococcus furiosus*, in complex with the active (methylated) AdoMet as well as with AdoHcy. The unusually high resolution of 1.2 Å allows for atomic level analysis of cofactor binding. Importantly, we also report the first structure of an enzyme of this type bound to a polypeptide substrate, VYP(L-isoAsp)HA. The various crystal structures implicate a significant protein conformational change during the catalytic cycle, and help explain how the enzyme achieves its critical specificity.

Results and Discussion

Multiple structures of the *Pyrococcus furiosus* L-isoaspartyl (D-aspartyl) methyltransferase

We purified a recombinant form of the L-isoaspartyl methyltransferase with six histidine residues added as an affinity tag to the C terminus (Figure 2). We determined crystal structures of the protein in binary complexes with different cofactors or cofactor analogs (specifically AdoMet, AdoHcy, and adenosine) and in a ternary complex with adenosine and a polypeptide substrate VYP(L-isoAsp)HA (Table 1). Initial phases were determined with a samarium derivative of a binary complex with the cofactor AdoHcy. The structure of this crystal form was resolved to the atomic resolution of 1.2 Å. The high-resolution atomic model obtained from the initial crystal form allowed for the straightforward refinement of the other structures to resolutions of 2.1 Å or better. While the VYP(L-isoAsp)HA polypeptide substrate was added to the protein sample during crystallization, the cofactors were all found serendipitously in various crystals grown from different preparations of the recombinant protein.

Protein fold and comparison to known structures

The crystal structures determined here for the L-isoaspartyl methyltransferase from *P. furiosus* show a Rossman-type fold as expected for a member of the superfamily of AdoMet-dependent methyltransferases (Figure 3).[28,29] Members of this family of enzymes almost always contain a central seven-stranded β-sheet connected by α-helices. The number of helices varies, but at least six α-helices in total separate the β-strands in the central sheet. The resulting overall fold is a doubly wound α/β/α sandwich structure.[29] The *P. furiosus* methyltransferase structure reported here reveals an unusual strand rearrangement at the edge of the core β-sheet compared to the typical sheet structure of this enzyme superfamily. The rearrangement can be described as an exchange of strands 6 and 7. This type of β-strand connectivity has only been observed to date in the L-isoaspartyl methyltransferase from *T. maritima*.[30] Our observation of this connectivity in a second L-isoaspartyl methyltransferase confirms the suggestion from amino acid sequence comparisons that L-isoaspartyl methyltransferase enzymes will all share this unique

Table 1. Data collection, phasing and refinement statistics

	Sm high remote	Sm inflection	Sm peak	AdoHyc	Adenosine	Adenosine with polypeptide	AdoMet
A. *Data collection*							
Wavelength (Å)	0.920	1.819	1.845	1.000	1.100	1.5418	1.100
Resolution limit (Å)	2.0	2.2	2.2	1.2	1.5	2.1	1.5
R_{sym}(%)	5.7	11.5	11.2	4.5	7.3	10.6	4.1
R_{sym} (%, last shell)	20.3	41.1	35.0	38.5	45.8	43.5	33.1
I/sigma (last shell)	5.8	1.8	4.1	2.6	2.4	2.9	6.6
Total observations	122,930	33,412	78,125	427,037	119,506	235,211	234,242
Unique reflections	25,518	10,455	10,755	61,904	30,345	39,722	33,412
Completeness (%)	96.2	97.3	99.9	97.7	95.3	99.6	99.6
Completeness (%, last shell)	83.3	95.6	99.7	86.6	97.8	99.9	99.9
B. *Phase determination*							
R_{cullis} (20-2.1 Å, acent/cent, iso)	0.75/0.75	N/A[a]	0.73/0.71				
R_{cullis} (20-2.1 Å, anom)	0.83	0.75	0.71				
Phasing power (20-2.1 Å, acent/cent)	1.28/0.97	N/A	1.33/0.99				
Mean overall figure of merit	0.61						
Mean overall figure of merit after dm	0.76						
C. *Model refinement*							
R_{free} (%, 20 Å, upper limit)				20.0	23.6	23.3	22.7
R_{work} (%, 20 Å, upper limit)				14.7	22.1	21.1	20.6
PDB ID code				1JG1	1JG2	1JG3	1JG4

	No. res.	Ave. B (Å^2)	r.m.s.d. Bonds (Å)	r.m.s.d. Angles (deg.)	B-values (Å^2 bonded)	
D. *Model quality-AdoHcy*					Main	Side
Protein	214	19.2	0.015	2.5	1.2	3.5
Water	256	38.4				
E. *Model quality-adenosine with polypeptide substrate*						
Protein	428	26.4	0.006	1.3	0.77	1.5
Water	225	35.2				

[a] The Sm inflection data set was treated as a reference for phasing calculations (see Materials and Methods).

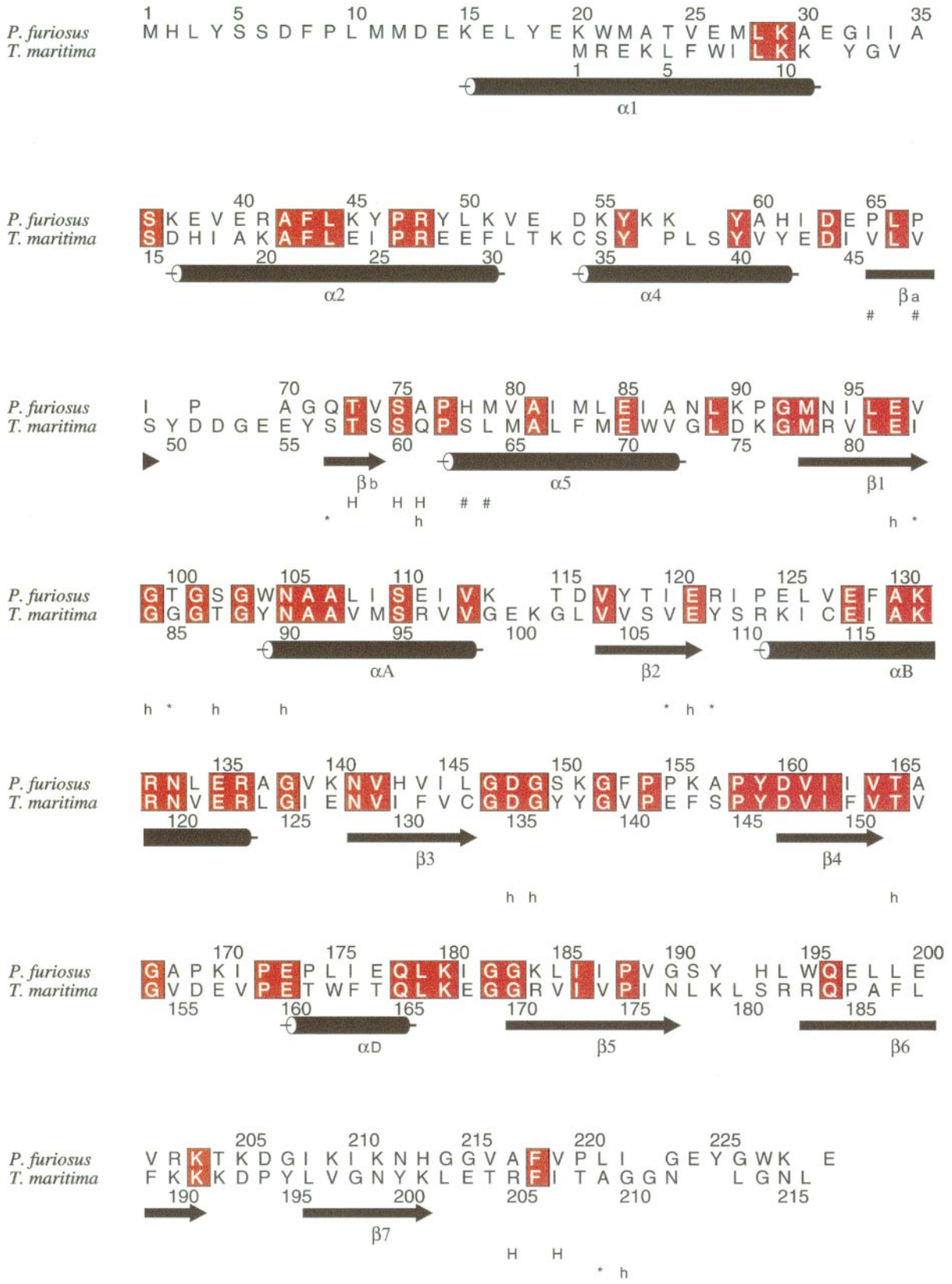


Figure 2. Structurally aligned sequences of the L-isoaspartyl (D-aspartyl) protein methyltransferases of *Pyrococcus furiosus* (this work) and *Thermotoga maritima*[18,30] by the program ALIGN_V2.[60] The *P. furiosus* amino acid sequence is taken from the genomic sequence MM1-MM1 02861 at the Utah Genomics Center (www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html). Conserved residues are shown in red. Secondary structure elements in the *P. furiosus* structure (this work) are shown using the naming scheme described by Skinner *et al*.[30] Capital H signifies hydrogen-bonding and # signifies hydrophobic interactions between the protein and the substrate polypeptide. Small h signifies hydrogen-bonding and an asterisk (*) signifies hydrophobic interaction between the protein and the AdoHcy cofactor. This Figure was made with ALSCRIPT.[61]

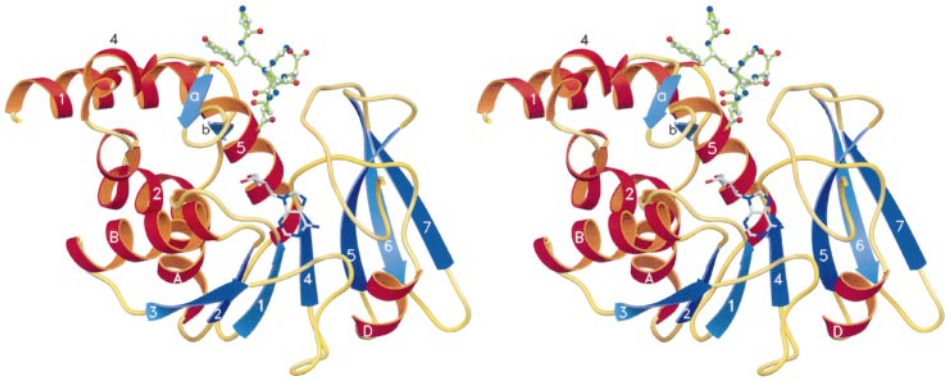


Figure 3. Stereo view of a ribbon diagram of the *P. furiosus* L-isoaspartyl methyltransferase. The structure shown is that of the ternary complex with adenosine in stick form and the peptide VYP(L-isoAsp)HA in ball-and-stick form. The secondary structures are labeled as in Figure 2. This Figure was made with SETOR.[62]

feature.[30] One possible evolutionary mechanism to explain the β-sheet rearrangement is a shortening of the long loop and helix between strands 5 and 6 (in the canonical fold), forcing β-strand 6 to lie adjacent to β-strand 5. Loop shortening has also been implicated in another form of structural rearrangement, namely "domain swapping" in which entire protein domains exchange positions.[31]

Overall, we find that the structure of the archaeal *P. furiosus* enzyme is very similar to that of the eubacterial *T. maritima* enzyme.[30] When the two proteins are aligned structurally, the root-mean-square deviation between corresponding backbone atoms is 1.4 Å (Figure 2).

Cofactor binding

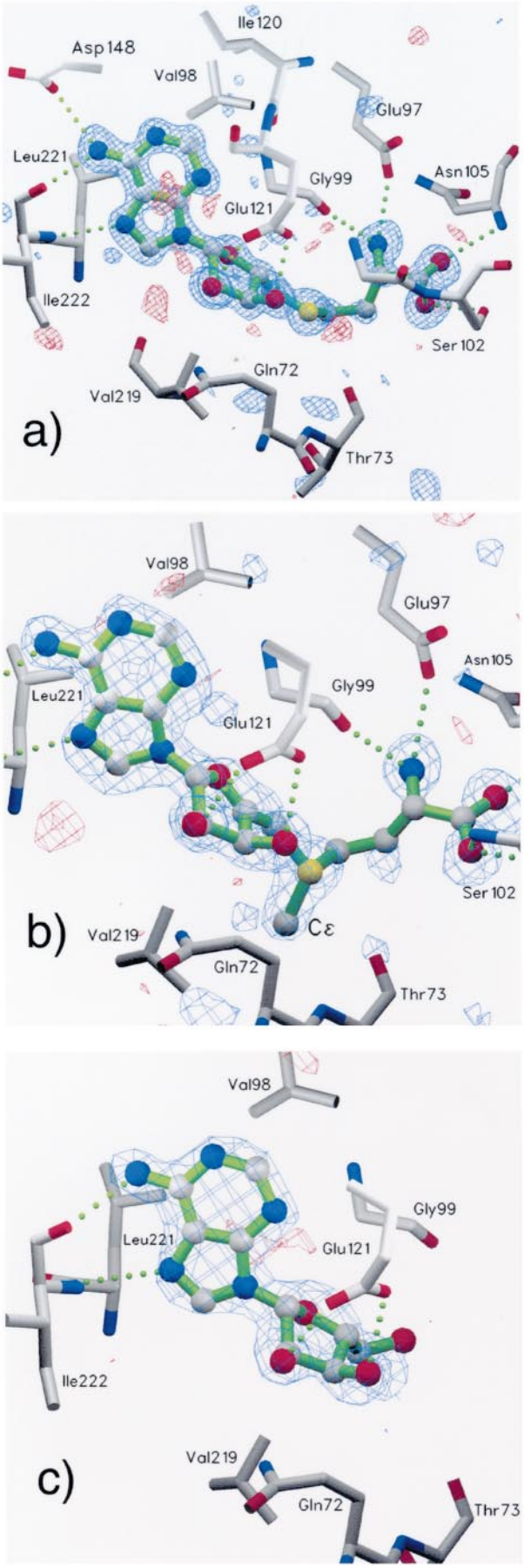
The quality of the initial diffraction data from the native and samarium-derivatized crystals was high enough that an automated model-building program was able to position nearly every atom of the known amino acid sequence into the experimental electron density map with high reliability.[32] The automated program also placed additional atoms into the structure that were readily recognized as the (unmethylated) cofactor, AdoHcy (Figure 4(a)). The occupancy of AdoHcy was approximately unity and its atomic displacement parameters were low for all atoms, except C^{γ} and C^{δ}. The presence of the AdoHcy cofactor was unexpected. No cofactor was added to either the protein preparation or the crystallization mixture. It thus appears that the cofactor was retained from the recombinant expression culture through a purification scheme that included an 80 °C heat treatment, elution from a nickel-chelating column with 300 mM imidazole, and dialysis for several days. The presence of stoichiometric (1:1) amounts of AdoHcy in the purified protein sample was verified by reverse-phase high performance chromatography using internal standards as described in Materials and Methods.

The AdoHcy molecule is located in a pocket of the protein, nearly completely buried, with approximately 1% of its surface accessible to solvent as calculated by the program AREAIMOL.[33] In Table 2, we compare the solvent exposure for AdoHcy and AdoMet cofactors found in other types of methyltransferases. A similar degree of cofactor burial is found in the catechol *O*-methyltransferase and the *cheR* protein L-glutamate methyltransferase, but the cofactors from other types of enzymes are much more exposed (Table 2). The affinity of mammalian isoaspartyl methyltransferases for AdoHcy is very high. For example, the K_i value for the bovine enzyme is 0.08 μM.[34] Furthermore, a dead-end complex of the mammalian methyltransferase with AdoHcy has been

Table 2. Comparison of cofactor exposure in various methyltransferases as calculated by AREAIMOL[33]

PDB id	Cofactor	Exposed (Å^2)
1JG1 L-isoaspartyl methyltransferase, *P. furiosus*	AdoHcy	8
1JG4 L-isoaspartyl methyltransferase, *P. furiosus*	AdoMet	5
1AF7 glutamate *O*-methyltransferase	AdoHcy	2
1VID catechol *O*-metyltransferase	AdoHcy	13
1DL5 L-isoaspartyl methyltransferase, *T. maritima*	AdoHcy	13 and 15
1XVA glycine *N*-methyltransferase	AdoMet	28
4MHT DNA C5-cytosine methyltransferase	AdoHcy	48
1BOO DNA N4-cytosine methyltransferase	AdoHcy	52
1V39 mRNA *O*-2′-methyltransferase	AdoHcy	65
2ADM DNA N6-adenine methyltransferase	AdoMet	107

[a] The total surface area of AdoHcy is 610 Å^2 and the total surface area of AdoMet is 618 Å^2.



detected,[34] which may correspond to a form of the enzyme in which cofactor exchange is greatly reduced. Such observations are consistent with the tendency of the *P. furiosus* enzyme to bind and retain various cofactors throughout purification. The atomic contacts between AdoHcy and the *P. furiosus* enzyme are shown in Figure 4(a) and the amino acid residues involved are indicated in Figure 2. The binding of AdoHcy in the present structure is similar to that reported recently for the L-isoaspartyl methyltransferase from *T. maritima*.[30]

Crystals with the same properties and unit cell as the AdoHcy crystals described above were grown from a second protein preparation. Refinement of the structural model against the diffraction data from these crystals indicated an almost complete loss of electron density for the sulfur atom and the amino acid region of the cofactor. The resulting cofactor electron density was best modeled by placing adenosine in the cofactor-binding site (Figure 4(c)). Two possible conformations for the 5′-hydroxyl group were observed to make good hydrogen bonding interactions with hydrogen bond acceptors in the protein. The protein preparation used to grow these crystals was verified to contain stoichiometric amounts of adenosine by HPLC as described in Materials and Methods.

The crystal structure of the adenosine binary complex of the *P. furiosus* L-isoaspartyl methyltransferase was not significantly different from that of the AdoHcy binary complex. The root-mean-square deviation of the backbone atoms was 0.14 Å. The significance of adenosine binding to the enzyme is not yet clear. However, adenosine is a central metabolite, and recent work has demonstrated that it can inhibit the activity of the *Pyrococcus furiosus* methyltransferase *in vitro* (N. T., S. G., T. O. Y. and S. C., unpublished results).

An additional type of binary complex was found in crystals prepared from a third protein preparation. The electron density maps showed clear density for the active methyl group (Figure 4(b)). The same result was found for three different crystals grown from the same protein preparation. The only significant difference observed while preparing this protein sample was a variation in the cell growth curve, which dropped unexpectedly from an $A_{600\ nm}$ of 2.1 to an $A_{600\ nm}$ of 1.1 during the third hour after induction. The presence of AdoMet was confirmed by HPLC, where crystals

Figure 4. Electron density surrounding the cofactors AdoHcy and AdoMet and the cofactor derivative adenosine in various structures of the *P. furiosus* L-isoaspartyl methyltransferase. (a) An $F_o - F_c$ electron density map around the AdoHcy cofactor at 1.2 Å resolution. (b) A simulated annealing omit map around the AdoMet cofactor at 1.5 Å resolution. (c) A simulated annealing omit map around the adenosine cofactor analog at 1.6 Å resolution. This Figure was made with SETOR.[62]

from the same crystallization tray were found to contain a 2:5 molar ratio of AdoMet to AdoHcy. The AdoHcy could be the result of spontaneous degradation of AdoMet over the four-month period between crystal growth and HPLC analysis. The electron density of the α, β, and γ carbon atoms of the methionine moiety were seen less clearly, although good density was observed for the α nitrogen and carboxyl oxygen atoms. As a result, we cannot rule out the partial occupancy of a degraded cofactor in these crystals. We considered the possibility that a spontaneous degradation product of AdoMet, 5′deoxy-5′-methylthioadenosine, might be present, but this compound could not be detected by HPLC analysis of the crystals.

Conformational changes correlated to the loss of the methyl group on the cofactor

The structure of the protein in the binary complex with AdoMet was similar but distinct from that of the binary complexes with AdoHcy (Figure 5). Histidine 193 points in towards the cofactor when a catalytic methyl group is present on the cofactor and away towards solvent when the catalytic group is absent. Tyrosine 192 also makes a dramatic change, swinging 10 Å and turning 90°. These changes result mainly from a ''peptide flip'' that occurs due to a 164° change in the backbone ϕ angle at residue 191 and a 100° change at residue 193. Despite this localized conformational difference, elsewhere the overall rms difference between the AdoHcy form of the protein and the AdoMet form is only 0.16 Å for the protein backbone atoms. Crystal contacts in this variable region include one hydrogen bond in each form, in addition to several less-specific interactions. The average atomic displacement parameter (*B*-factor) for the loop is 27 $Å^2$ for the AdoMet configuration and 24 $Å^2$ for the AdoHcy configuration.

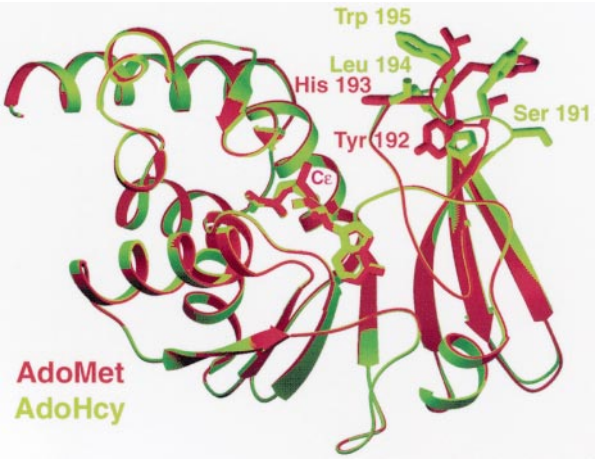


Figure 5. A protein conformational difference between the AdoMet (red) and AdoHcy (green) complexes of the *P. furiosus* isoaspartyl methyltransferase. The largest coordinate differences are over 10 Å. This Figure was produced with SETOR.[62]

The observed conformational change may be important for the activity of the enzyme. Structures refined against independent data sets from three different AdoMet-containing crystals all showed the same significant structural rearrangement in the loop between residues 190 and 195 (Figure 5). Because of this correlation, we suggest that the observed structural difference reflects a real conformational change that occurs during the catalytic cycle. This change could represent part of a larger conformational change that would have to occur for the cofactor to exchange, or it could be an integral part of the catalytic mechanism, or it could play a combination of these exchange and catalytic roles. For instance, binding of a substrate to the protein could energetically favor the conformation of the protein that preferentially binds AdoHcy over the conformation that preferentially binds AdoMet. This would allow substrate binding to facilitate the transfer of the methyl group from AdoMet to the substrate.

The protein conformational change occurs near the cofactor, with the closest atom of the loop (His193 $N^{\varepsilon 2}$) being approximately 7 Å from the catalytic methyl group. The catalytic methyl group is likely the key atom in affecting the observed protein conformational change. Two potential triggers could be identified near the catalytic methyl group, which might initiate the conformational change observed in loop 190-195. First, valine 219 is within van der Waal's distance to the catalytic methyl group (3.5 Å). In the absence of the methyl group, valine 219 moves (0.13 Å) closer to the position vacated by the catalytic methyl group. Second, a water molecule occupies the physical space vacated by the catalytic methyl in all of the structures where that group is missing. This water molecule makes three hydrogen bonds with the side-chain and backbone atoms of threonine 73 and a hydrogen bond to a second water molecule. These hydrogen bonds are lost when AdoMet is bound, providing a potential energetic connection between the methylation state of the cofactor and the conformation of the protein near the active site.

Substrate binding and specificity

L-Isoaspartyl (D-aspartyl) methyltransferases have been found to methylate only two of the four possible aspartyl-based amino acids, namely L-isoaspartyl and D-aspartyl residues. No reactivity has been seen with L-aspartyl or D-isoaspartyl residues (Figure 6).[35] Discrimination against methylation of natural L-aspartyl residues is crucial to the function of the enzyme; methylation of the natural amino acid would actively generate isomerized and racemized residues by increasing the steady-state concentration of the succinimidyl intermediate. A possible model to explain the enzyme's substrate specificity is shown in Figure 6, where the four aspartyl-based isomers are drawn in extended

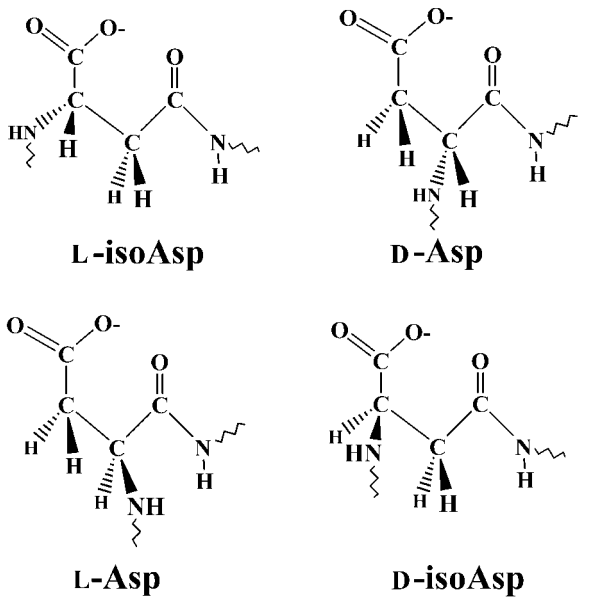


Figure 6. Covalent structures of the isomerized and racemized aspartyl residues that occur in damaged proteins. Only the L-isoaspartyl and D-aspartyl forms shown at the top are substrates for the methyltransferase. The enzyme is able to catalyze a net conversion of damaged residues to natural L-aspartyl residue by discriminating between different isomers, specifically by failing to act on the natural L-aspartate.

configurations with their potentially reactive carboxyl groups positioned similarly. The two species that are not substrates both have their amide nitrogen atoms facing out of the page, while the two species that are substrates both have their nitrogen atoms facing into the page. The amide nitrogen is the atom to which the N-terminal part of the polypeptide substrate must be connected. The model therefore suggests how steric effects could distinguish the two recognizable substrates from the two that are unrecognizable.

Prior to determining the crystal structure of the enzyme-substrate complex, we undertook computer-based substrate docking studies to address the issue of substrate specificity in more detail. It was possible to fit a polypeptide containing an L-isoaspartyl residue into the narrow active site (of the AdoMet binary complex structure) by a combination of manual and computer-based methods. First, we manually placed the α-carboxyl group of the substrate in a position where it could attack the methyl group on the cofactor. By visual inspection we found that the L-isoaspartate carboxylate group could be docked near the methyl group of the AdoMet donor while making geometrically acceptable hydrogen bonds with the backbone nitrogen atom of Val219 and the hydroxyl group of Ser75, both conserved residues in L-isoaspartyl methyltransferases from various organisms. Having placed the L-isoaspartyl residue in the active site, the C-terminal portion of the polypeptide substrate L-isoAsp-Gly-Phe could be modeled into a cleft running alongside the active site. The program AUTODOCK was then used to optimize the conformation of the peptide.[36] In the resulting model (not shown), several energetically stabilizing interactions were found. Also, the amide nitrogen atom of the L-isoaspartyl residue was found to point out into the solvent space, as expected.

Following the modeling studies, the enzyme was successfully co-crystallized with a cognate polypeptide, VYP(L-isoAsp)HA and the structure was determined to a resolution of 2.1 Å. These adenosine and polypeptide substrate ternary complex crystals grew in a space group different from that of the adenosine complex. Despite potential differences arising from the substrate and from distinct crystal packing, only small differences were observed between the two protein structures. The rms difference was only 0.43 Å over backbone atoms. The small differences in backbone positions appear to result from a small hinge-type movement that causes the protein to close slightly upon the active site. The crystal structure of the substrate in the enzyme active site is shown in Figure 7. The polypeptide substrate binds in an essentially extended conformation. The majority of the stabilizing interactions come from hydrogen-bonding to the polypeptide backbone of the enzyme (Figure 8). The main features of the ternary complex are consistent with the model predicted in advance by manual and computer-aided docking. Five hydrogen bonds were correctly predicted between the substrate and the protein. Comparing the docking model to the experimental structure of the substrate, an rms difference of 0.99 Å was calculated over 12 atoms including the cognate isoaspartyl

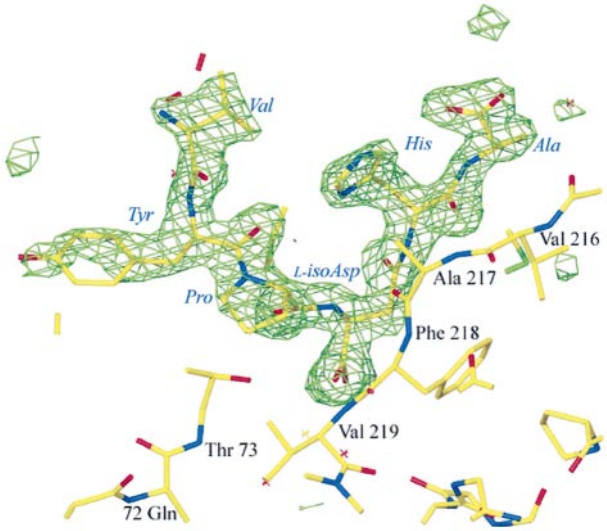


Figure 7. An isoaspartyl polypeptide co-crystallized with the isoaspartyl methyltransferase repair enzyme. A simulated annealing omit map is shown in the region of the bound substrate at 2.1 Å resolution. This Figure was made with the program O.[56]

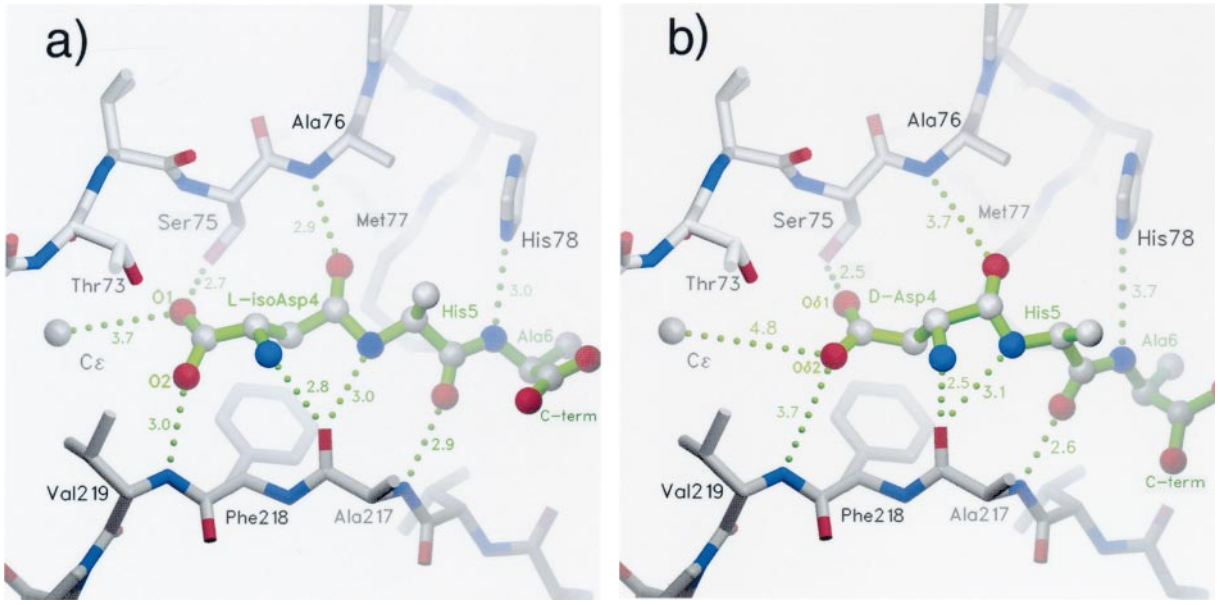


Figure 8. The binding of peptide substrates to L-isoaspartyl methyltransferase. The experimental structure is shown (left) of the VYP(L-isoAsp)HA substrate bound to the enzyme (as a ternary complex with adenosine). The region of the substrate N-terminal to the reactive isoaspartyl would emerge out of the page and is not shown. The position that would be occupied by the methyl group of the AdoMet cofactor is shown by the grey sphere labeled C^{ε}. A theoretical structure is also shown (right) which was produced by automatic docking of the similar polypeptide containing a D-aspartyl residue instead of the L-isoaspartyl residue. Similar models with L-aspartate and D-isoaspartate could not be generated without severe steric clashes, consistent with the known specificity of the enzyme. This Figure was made with SETOR.[62]

residue and part of the following histidine residue. The general agreement between the experimental result and the computer modeling result verified that the docking approach can be reliable in situations like the present one in which there is an identifiable reactive group in a deep active site cleft. The agreement also justified using the docking approach in subsequent analyses (below).

As discussed above, it is critical for the L-isoaspartyl methyltransferase to be able to discriminate between natural and unnatural isomers of aspartyl residues. The crystal structure of the enzyme-substrate complex serves as a framework for understanding this discrimination. The structure of the complex with the L-isoaspartyl-containing polypeptide makes it reasonably clear how the D-aspartyl residue can be accommodated and L-aspartyl and D-isoaspartyl residues can be excluded. The enzyme binds the substrate primarily through interactions with the isoaspartyl residue and the region of the polypeptide C-terminal to that residue. This justifies the earlier analysis[10] (Figure 6), in which it was pointed out that if the side-chain carboxylate group and the C-terminal polypeptide backbone are held in place, then the N terminus of the substrate emerges from one side of the peptide plane for the two residue types that can be recognized (L-isoaspartyl and D-aspartyl), while it emerges from the other side of the peptide plane for the two residue types that cannot be recognized (D-isoaspartyl and natural L-aspartyl).

For the permitted substrates, the N terminus of the polypeptide protrudes out into the solvent. For the forbidden substrates, the polypeptide N-terminus makes severe steric clashes with the enzyme. Substrates containing D-isoaspartate collide at residue Pro65 when their C terminus is modeled in the position found in the substrate crystal structure. The collision occurs at Phe218 if the substrate contains L-aspartate.

In contrast, it is relatively easy to fit the corresponding D-aspartyl polypeptide into the enzyme active site. After manually placing the nucleophilic carboxylate of the model substrate adjacent to the reactive methyl position, the program AUTODOCK was run as before to optimize the binding energy and configuration.[36] The D-aspartyl polypeptide was able to maintain five hydrogen bonds to the enzyme. However, three hydrogen bonds were stretched from approximately 3.0 Å to 3.7 Å. This is beyond the limit of a hydrogen bond, perhaps explaining the poorer affinity for D-aspartyl *versus* L-isoaspartyl residues. Furthermore, the reactive carboxylate oxygen atom fell 1.1 Å further from the reactive methyl group, potentially reducing the enzymatic activity towards D-aspartyl residues. However, because these differences depend on the reliability of the modeling results, experimental results would be required to verify the predictions.

Focusing on the conformation of the D-aspartyl polypeptide modeled into the active site, a second

unexpected mechanism is revealed by which the enzyme seems to discriminate against polypeptides containing natural L-aspartyl residues. The enzyme binds the D-aspartyl polypeptide in such a way that its backbone torsion angles would be incompatible with the presence of an L-amino acid. When the configuration of the bound D-aspartyl substrate is optimized by AUTODOCK[36] the ϕ and ψ torsion angles at the D-aspartyl residue are found to be 121° and 43°. A Ramachandran diagram (which must be turned upside down for D-amino acids) shows that the observed torsion angles are energetically allowed for a polypeptide containing a D-amino acid, but not a natural L-amino acid (Figure 9). Apparently, a polypeptide containing an L-aspartyl residue cannot adopt the backbone conformation required for binding to the enzyme without experiencing internal collisions between its own atoms.

Structural phylogeny in the extended methyltransferase family

With over a dozen structures of enzymes from the AdoMet-dependent methyltransferase superfamily now available, it is possible to analyze their evolutionary relationships. Utilizing the program Multiple Alignment of Protein Structures (MAPS),[37] 12 methyltransferase structures were aligned to produce a structural phylogeny (Figure 10). Interestingly, we find that the protein L-isoaspartyl methyltransferase is structurally more closely related to the small molecule glycine *N*-methyltransferase[38] than to the *cheR* protein L-glutamyl *O*-methyltransferase,[39,40] which also methyl esterifies a carboxyl group. From a wider view, a branch containing protein and small molecule methyltransferases (including the catechol *O*-methyltransferase[41] and the RMT1[42] and PRMT3[43] protein arginine *N*-methyltransferases) can be clearly distinguished from a branch containing the *C*-, *N*-, and *O*- DNA and RNA methyltransferases (Figure 10). These findings suggest an early divergence of the nucleic acid and protein/small molecule methyltransferases and the possible evolution of the protein L-isoaspartyl methyltransferase from a small molecule methyltransferase.

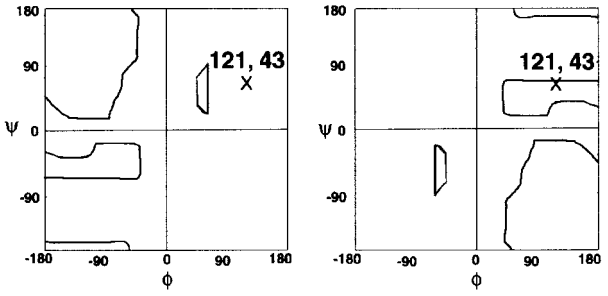


Figure 9. Ramachandran plots showing the allowed torsion angles for a polypeptide backbone. The allowed angles for a (non-glycine) natural L-amino acid are shown on the left. For a D-amino acid (right) there is a negation of the allowed angles, so the plot must be turned upside down. In a model of a polypeptide substrate containing a D-aspartyl residue bound in the active site of the repair enzyme, the torsion angles of the backbone at the reactive residue are $\phi = 121°$, $\psi = 43°$. The plots show that this local backbone conformation can be adopted by a D-amino acid but not by a natural L-amino acid. This intramolecular steric barrier helps explain why a polypeptide containing only L-amino acids cannot adopt the conformation necessary to bind in the active site of this repair enzyme.

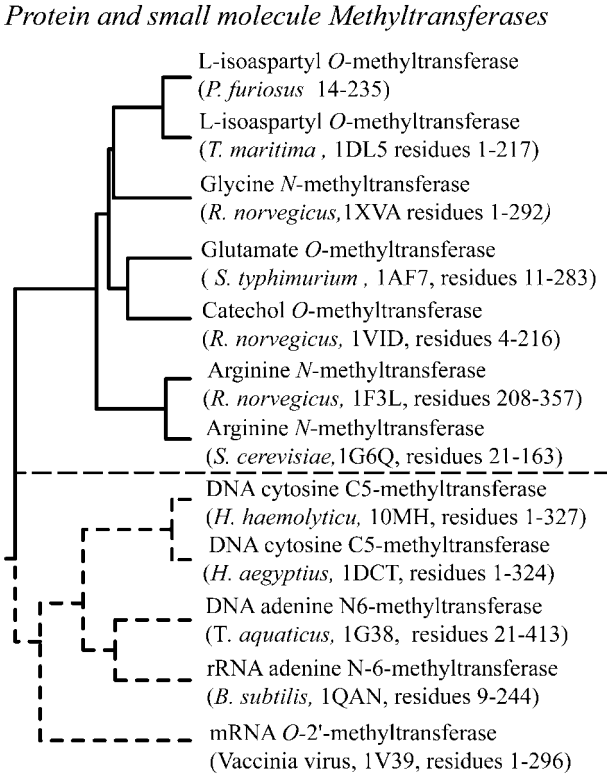


Figure 10. Structural phylogeny of AdoMet-methyltransferases derived from multiple aligned protein structures.[37] The lengths of the horizontal bars indicate the relative structural differences between the methyltransferases. The structural comparisons automatically separate the methyltransferases that act on small molecules or isoaspartyl residues in proteins from those that operate on nucleic acids. The family of protein and small molecule methyltransferases is shown with continuous lines; the family of nucleic acid methyltransferases is shown with dotted lines. The residues used for comparison are indicated in each case, along with the RCSB PDB accession code in parentheses.

Conclusions

The *P. furiosus* L-isoaspartyl methyltransferase has a structure typical of the AdoMet-dependent methyltransferase superfamily, except for an exchange of two strands at the edge of the central β-sheet. This unusual feature appears to be charac-

teristic of the L-isoaspartyl methyltransferase enzymes. A comparison of the many known methyltransferase structures shows that L-isoaspartyl methyltransferase resembles more closely the small molecule methyltransferases than the nucleic acid methyltransferases.

Multiple structures of the *P. furiosus* enzyme from various crystals show that the enzyme binds to various forms and analogs of its AdoMet cofactor tightly enough to retain them throughout purification. The presence or absence of the catalytic methyl group on the cofactor correlates strongly with a dramatic conformational change in loop 190-195. Because the bound cofactor is converted from AdoMet to AdoHcy during the catalytic cycle, the observed conformational change could be important in catalysis. The nearly complete burial of the cofactor in the enzyme active site also argues that such protein conformational changes would be required for cofactor exchange.

The structure of a ternary complex shows for the first time how L-isoaspartyl methyltransferase recognizes and methylates damaged polypeptide substrates. The enzyme appears to use two mechanisms to make the critical distinction between damaged and natural isomers of aspartyl residues in proteins. First, a polypeptide containing a natural L-aspartyl residue cannot bind in the deep active site without colliding with the enzyme. Second, the backbone of a polypeptide which contains only natural L-amino acids, cannot adopt the torsion angles necessary for binding without experiencing internal steric clashes. This latter strategy represents a new element in the set of diverse mechanisms enzymes use to recognize and repair molecular damage.

Materials and Methods

Cloning, expression, purification

The *P. furiosus* methyltransferase gene was identified with a psi-BLAST search of the data from the Utah Genome Center as deposited with the NCBI unfinished microbe genomes database (www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html).[44] We PCR-amplified the gene (MM1-MM1 02861) using a 5′ primer corresponding to the sequence immediately upstream from the first AUG codon in the reading frame (5′-ATTATCGAAAGCACGACTACAAGG-3′) and a 3′ complementary primer that replaced the original stop codon with six histidine codons followed by a stop codon (5′-CTAATGATGATGATGATGATGTTCTTTCCACCCGTATTCCCC-3′). The template DNA was provided as a BAC-clone prepared from *P. furiosus* genomic DNA and was a generous gift of the Utah Genome Center in Salt Lake City, UT. The PCR product was cloned directly into the pTrcHis2-TOPO vector in TOP10 *Escherichia coli* cells by topo-isomerase-mediated cloning (Invitrogen, San Diego, CA).[45] An ampicillin-resistant colony was selected that gave good expression in crude extracts of a 25 kDa polypeptide species by SDS-PAGE when grown to an $A_{600\ nm}$ of 0.7 at 37 °C and induced with 1 mM isopropyl thio-β-D-galactoside (IPTG) and grown for an additional two hours. Methyltransferase was produced by initially inoculating one litre of LB broth with 100 μg of ampicillin and 2 % (w/v) glucose. After growing overnight at 37 °C, a portion was used to inoculate ten litres of enriched growth media to an initial A_{600} of approximately 0.1. Cells were grown at 37 °C. When the cell density reached an $A_{600\ nm}$ of 0.8, IPTG was added to a final concentration of 1 mM and the incubation continued for three hours. Cells (3-10 g wet weight) were harvested by centrifugation, resuspended in 20 ml/g of 0.5 M NaCl, 50 mM Tris-HCl (pH 7.0), 1 mM benzamadine and 5 μM leupeptin. Cells were disrupted by incubation for 20-30 minutes at room temperature with 3 mg of hen egg-white lysozyme and 2 mg of bovine pancreatic DNase I, followed by sonication on ice for a total of six minutes and two passes through a French press at 10,000 lb/in^2 (1 lb/in$^2 \approx 6.9$ kPa). The soluble fraction (100-200 ml) was separated from the pellet and from loosely adhering material by centrifugation at 20,000 ***g*** at 5 °C for 45-60 minutes. This fraction was brought to 80 °C over a 15 minute time period in a 250 ml Erlenmeyer flask with occasional stirring and was then maintained at 75-80 °C for 15 minutes. After incubation on ice for 15 minutes, the opaque solution was centrifuged as before. The soluble fraction was then loaded onto a 5 ml nickel-containing Hitrap column (Amersham Pharmacia) and eluted with 125 ml of a 0-0.5 M gradient of imidazole-HCl in Hepes (pH 7.8), 350 mM NaCl. Fractions (1 ml) were collected at a flow rate of 1 ml/minute. Fractions eluting at 25-50 ml were found to contain the 25 kDa polypeptide at an approximate purity of 98 % by SDS-PAGE. These fractions were concentrated and dialyzed into 0.5 M NaCl, 50 mM Tris-HCl (pH 7.0) using a Prodicon apparatus with 10 kDa filter to give a final concentration of 12 mg/ml. Protein samples were obtained from several independent preparations. Although the same protocol was followed, significant differences were observed in cell growth curves. N-terminal Edman analysis was performed by Midwest Analytical, Inc. (St. Louis, MO) and gave a sequence MHLYS. This result indicates that translation in *E. coli* apparently starts at a methionine residue 43 codons downstream of the first potentially encoded methionine residue in the open reading frame to give a product of 235 residues. The protein had a specific activity at pH 7.5 of 12,000 pmol methyl groups transferred/minute per mg protein at 70 °C and 27,000 pmol/minute per mg protein at 85 °C using VYP(L-isoAsp)HA as a methyl-acceptor according to the assay described by Gilbert and others.[46]

Crystal structure determination

Native crystals of the *P. furiosus* methyltransferase grew in two to seven days at 24 °C by the hanging drop method. A 1 μl volume of 12 mg/ml protein in 0.5 M NaCl, 20 mM Hepes (pH 7.0) buffer was mixed with 1 μl of a well solution containing 25 % (w/v) polyethylene glycol 2000, 15 % (v/v) glycerol, 0.2 M sodium acetate (pH 4.6). The mixture was then allowed to equilibrate over the well solution. The crystals belong to space group $P2_12_12_1$ ($a = 39.2$ Å, $b = 52.5$ Å, $c = 96.6$ Å). Candidates for heavy-atom derivatives were screened by the gel shift method of Boggon & Shapiro.[47] Protein was incubated for ten minutes at room temperature in 28 different 1 mM solutions of heavy metals and analyzed by native PAGE. A clear gel shift was found for samarium chloride. Native methyltransferase crystals were briefly soaked in samarium tetrachloride, producing a derivative suitable for multiwavelength anomalous dispersion (MAD) phase determination.

Diffraction data (for structures with the PDB codes 1jg1, 1jg2 and 1jg4) were collected at Brookhaven National Laboratory using beamline X8C of the National Synchrotron Light Source and reduced using DENZO/SCALEPACK.[48] A single samarium derivative crystal was used to collect MAD data sets at three wavelengths, with the 1.819 Å inflection wavelength treated as a reference using the program SOLVE.[49] Three samarium sites were found in the anomalous difference Patterson maps using SHELXD.[50] After phasing with MLPHARE[33] and applying density modification with DM[33,51] a figure of merit of 0.76 was achieved and an interpretable density map at 2.1 Å was calculated with CNS.[52] The high quality of the native data set combined with the MAD phases permitted the program ARP (Automatic Refinement Program) to trace nearly the entire protein chain.[32] A few rounds of refinement with SHELXL[53] and the addition of 256 well-ordered water molecules led to a final structure with a working *R*-value of 14.7 % and a free *R* value of 20.0 %. The final structure was validated with Ramachandran plots and the programs ProCheck[54] and Errat[55]. The AdoHcy structure was used as a starting model for refinement of the AdoMet and adenosine structures using CNS.[52] Refinement continued with CNS[52] for several rounds with model adjustments in O.[56] Final statistics are given in Table 1.

For crystals containing the methyl-accepting peptide substrate, a tenfold molar excess of VYP(L-isoAsp)HA (California Peptide Research, Napa, CA; Lowenson & Clarke[57]) was added to the protein solution and then mixed with an equal volume of 50 % (v/v) 2-methyl-2,4-pentanediol, 100 mM Tris-HCl, 200 mM ammonium phosphate (pH 8.5) (Hampton Crystal Screen 2, number 43). Crystals grew in three days at 22 °C by vapor diffusion and belong to space group $P6_1$ ($a = b = 91.6$ Å, $c = 124.6$ Å). Diffraction data were collected with a Rigaku FRD rotating copper anode generator at UCLA and processed with DENZO/SCALEPACK.[48] The AdoHcy structure was used as a search model for molecular replacement in the program EPMR.[58] The structure of the ternary complex (PDB code 1jg3) was refined using CNS.[52]

Chromatographic analysis of cofactors

The presence and identity of the cofactors in different protein preparations were confirmed by reverse-phase HPLC analysis. Protein samples (either 1 µl of liquid or one crystal) were incubated in 25 µl of 4 % (w/v) perchloric acid at room temperature for 30 minutes, centrifuged for three minutes at 20,800 *g* and the supernatant was then analyzed for adenosine derivatives using a modified protocol.[59] Briefly, the sample was chromatographed using an Alltech Econosphere C18 column (5 µm pore size, 250 mm long, 4.6 mm inner diameter) with a matching 7.5 mm long guard column. The column was eluted with a gradient of buffer A (50 mM sodium phosphate (pH 3.2), 10 mM heptane sulfonic acid, 4 % (v/v) acetonitrile) to buffer B (100 % acetonitrile) linearly from 100 % to 84 % buffer A over 16 minutes, 84 % to 79 % over nine minutes, then 79 % to 5 % over three minutes. The column was finally held at 5 % buffer A for six minutes. All column runs were preceded by a 20 minute equilibration in buffer A. Absorbance was measured at 254 nm. We found that AdoHcy eluted at 18 minutes, AdoMet eluted at 19.5 minutes and adenosine eluted at 16 minutes.

Accession numbers

The structures presented here are available from the RCSB PDB with their corresponding structure factors under accession numbers 1jg1 for the AdoHcy form, 1jg2 for the adenosine form, 1jg3 for the ternary adenosine form with L-isoaspartyl polypeptide substrate, and 1jg4 for the AdoMet form.

Acknowledgments

The authors thank Dr Duilio Cascio and Cameron Mura for useful discussions. We also thank the members of the PRT for Brookhaven National Lab beamline X8C. This work was supported by USPHS Predoctoral Training Program GMO7185 (to S.C.G. and J.E.K.), Department of Energy grant DE-FC03-87ER60615 (to T.O.Y.), and NIH grants AG18000 and GM26020 (to S.C.).

References

1. Geiger, T. & Clarke, S. (1987). Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* **262**, 785-794.
2. Bhatt, N. P., Patel, K. & Borchardt, R. T. (1990). Chemical pathways of peptide degradation. I. Deamidation of adrenocorticotropic hormone. *Pharm. Res.* **7**, 593-599.
3. Radkiewicz, J. L., Zipse, H., Clarke, S. & Houk, K. N. (2001). Neighboring side-chain effects on asparaginyl and aspartyl degradation: an *ab initio* study of the relationship between peptide conformation and backbone NH acidity. *J. Am. Chem. Soc.* **135**, 3499-3606.
4. Aswad, D. W., Paranandi, M. V. & Schurter, B. T. (2000). Isoaspartate in peptides and proteins: formation, significance, and analysis. *J. Pharm. Biomed. Anal.* **21**, 1129-1136.
5. Najbauer, J., Orpiszewski, J. & Aswad, D. W. (1996). Molecular aging of tubulin: accumulation of isoaspartyl sites *in vitro* and *in vivo*. *Biochemistry,* **35**, 5183-5190.
6. Tarcsa, E., Szymanska, G., Lecker, S., O'Connor, C. M. & Goldberg, A. L. (2000). Ca^{2+}-free calmodulin and calmodulin damaged by *in vitro* aging are selectively degraded by 26 S proteasomes without ubiquitination. *J. Biol. Chem.* **275**, 20295-20301.
7. Esposito, L., Vitagliano, L., Sica, F., Sorrentino, G., Zagari, A. & Mazzarella, L. (2000). The ultrahigh resolution crystal structure of ribonuclease A containing an isoaspartyl residue: hydration and sterochemical analysis. *J. Mol. Biol.* **297**, 713-732.
8. Mamula, M. J., Gee, R. J., Elliott, J. I., Sette, A., Southwood, S., Jones, P. J. & Blier, P. R. (1999). Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. *J. Biol. Chem.* **274**, 22321-22327.
9. Szymanska, G., Leszyk, J. D. & O'Connor, C. M. (1998). Carboxyl methylation of deamidated calmodulin increases its stability in *Xenopus oocyte* cytoplasm. Implications for protein repair. *J. Biol. Chem.* **273**, 28516-28523.
10. Clarke, S. (1999). A protein carboxyl methyltransfease that recognizes and repairs age-damaged pep-

tides and proteins and participates in their repair. In *S-Adenosylmethionine-dependent Methyltransferases: Structures and Functions* (Cheng, X. & Blumenthal, R. M., eds), pp. 123-148, World Scientific Pub,. River Edge, NJ.
11. Johnson, B. A., Murray, E. D., Jr., Clarke, S., Glass, D. B. & Aswad, D. W. (1987). Protein carboxyl methyltransferase facilitates conversion of atypical L-isoaspartyl peptides to normal L-aspartyl peptides. *J. Biol. Chem.* **262**, 5622-5629.
12. McFadden, P. N. & Clarke, S. (1987). Conversion of isoaspartyl peptides to normal peptides: implications for the cellular repair of damaged proteins. *Proc. Natl Acad. Sci. USA,* **84**, 2595-2599.
13. Brennan, T. V., Anderson, J. W., Jia, Z., Waygood, E. B. & Clarke, S. (1994). Repair of spontaneously deamidated HPr phosphocarrier protein catalyzed by the L-isoaspartate (D-aspartate) *O*-methyltransferase. *J. Biol. Chem.* **269**, 24586-24595.
14. Radkiewicz, J. L., Zipse, H., Clarke, S. & Houk, K. N. (1996). Accelerated racemization of aspartic acid and asparagine residues via succinimide intermediates: An *ab initio* theoretical exploration of mechanism. *J. Am. Chem. Soc.* **118**, 9148-9155.
15. Lowenson, J. D. & Clarke, S. (1992). Recognition of D-aspartyl residues in polypeptides by the erythrocyte L- isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. *J. Biol. Chem.* **267**, 5985-5995.
16. Kagan, R. M., McFadden, H. J., McFadden, P. N., O'Connor, C. & Clarke, S. (1997). Molecular phylogenetics of a protein repair methyltransferase. *Comp. Biochem. Physiol. sect B,* **117**, 379-385.
17. Mudgett, M. B., Lowenson, J. D. & Clarke, S. (1997). Protein repair L-isoaspartyl methyltransferase in plants. Phylogenetic distribution and the accumulation of substrate proteins in aged barley seeds. *Plant Physiol.* **115**, 1481-1489.
18. Ichikawa, J. K. & Clarke, S. (1998). A highly active protein repair enzyme from an extreme thermophile: The L-isoaspartyl methyltransferase from *Thermotoga maritima*. *Arch. Biochem. Biophys.* **358**, 222-231.
19. Visick, J. E., Cai, H. & Clarke, S. (1998). The L-isoaspartyl protein repair methyltransferase enhances survival of aging *Escherichia coli* subjected to secondary environmental stresses. *J. Bacteriol.* **180**, 2623-2629.
20. Kim, E., Lowenson, J. D., Maclaren, D. C., Clarke, S. & Young, S. G. (1997). Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice. *Proc. Natl Acad. Sci. USA,* **94**, 6132-6137.
21. Kagan, R. M., Niewmierzycka, A. & Clarke, S. (1997). Targeted gene disruption of the *Caenorhabditis elegans* L-isoaspartyl protein repair methyltransferase impairs survival of dauer stage nematodes. *Arch. Biochem. Biophys.* **348**, 320-328.
22. Yamamoto, A., Takagi, H., Kitamura, D., Tatsuoka, H., Nakano, H., Kawano, H. *et al*. (1998). Deficiency in protein L-isoaspartyl methyltransferase results in a fatal progressive epilepsy. *J. Neurosci.* **18**, 2063-2074.
23. Kim, E., Lowenson, J. D., Clarke, S. & Young, S. G. (1999). Phenotypic analysis of seizure-prone mice lacking L-isoaspartate (D-aspartate) *O*-methyltransferase. *J. Biol. Chem.* **274**, 20671-20678.
24. Lowenson, J. D., Kim, E., Young, S. G. & Clarke, S. (2001). Limited accumulation of damaged proteins in L-isoaspartyl (D-aspartyl) *O*-methyltransferase-deficient mice. *J. Biol. Chem.* **276**, 20695-20702.
25. McFadden, P. N. & Clarke, S. (1982). Methylation at D-aspartyl residues in erythrocytes: possible step in the repair of aged membrane proteins. *Proc. Natl Acad. Sci. USA,* **79**, 2460-2464.
26. Ingrosso, D., Fowler, A. V., Bleibaum, J. & Clarke, S. (1989). Sequence of the D-aspartyl/L-isoaspartyl protein methyltransferase from human erythrocytes. Common sequence motifs for protein, DNA, RNA, and small molecule *S*-adenosylmethionine-dependent methyltransferases. *J. Biol. Chem.* **264**, 20131-20139.
27. Kagan, R. M. & Clarke, S. (1994). Widespread occurrence of three sequence motifs in diverse *S*-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch. Biochem. Biophys.* **310**, 417-427.
28. Schluckebier, G., O'Gara, M., Saenger, W. & Cheng, X. (1995). Universal catalytic domain structure of AdoMet-dependent methyltransferases. *J. Mol. Biol.* **247**, 16-20.
29. Fauman, E. B., Blumenthal, R. M. & Cheng, X. (1999). Structure and evolution of AdoMet-dependent methyltransferase. In S-*Adenosylmethionine-dependent Methyltransferases: Structures and Functions* (Cheng, X. & Blumenthal, R. M., eds), pp. 1-38, World Scientific, River Edge, NJ.
30. Skinner, M. M., Puvathingal, J. M., Walter, R. L. & Friedman, A. M. (2000). Crystal structure of protein isoaspartyl methyltransferase: a catalyst for protein repair. *Structure,* **8**, 1189-1201.
31. Bennett, M. J., Choe, S. & Eisenberg, D. (1994). Domain swapping: entangling alliances between proteins. *Proc. Natl Acad. Sci. USA,* **91**, 3127-3131.
32. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458-463.
33. Collaborative Computational Project No. 4 (1994). CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.
34. Johnson, B. A. & Aswad, D. W. (1993). Kinetic properties of bovine brain protein L-isoaspartyl methyltransferase determined using a synthetic isoaspartyl peptide substrate. *Neurochem. Res.* **18**, 87-94.
35. Murray, E. D., Jr & Clarke, S. (1984). Synthetic peptide substrates for the erythrocyte protein carboxyl methyltransferase. Detection of a new site of methylation at isomerized L-aspartyl residues. *J. Biol. Chem.* **259**, 10722-10732.
36. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. & Olson, A. J. (1998). Automated docking using Lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **19**, 1639-1662.
37. Lu, G. (2000). TOP: a new method for protein structure comparisons and similarity searches. *J. Appl. Crystallog.* **33**, 176-183.
38. Fu, Z., Hu, Y., Konishi, K., Takata, Y., Ogawa, H., Gomi, T., Fujioka, M. & Takusagawa, F. (1996). Crystal structure of glycine *N*-methyltransferase from rat liver. *Biochemistry,* **35**, 11985-11993.
39. West, A. H., Martinez-Hackert, E. & Stock, A. M. (1995). Crystal structure of the catalytic domain of the chemotaxis receptor methylesterase, CheB. *J. Mol. Biol.* **250**, 276-290.
40. Djordjevic, S. & Stock, A. M. (1997). Crystal structure of the chemotaxis receptor methyltransferase

CheR suggests a conserved structural motif for binding *S*-adenosylmethionine. *Structure,* **5**, 545-558.
41. Vidgren, J., Svensson, L. A. & Liljas, A. (1994). Crystal structure of catechol *O*-methyltransferase. *Nature,* **368**, 354-358.
42. Weiss, V. H., McBride, A. E., Soriano, M. A., Filman, D. J., Silver, P. A. & Hogle, J. M. (2000). The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nature Struct. Biol.* **7**, 1165-1171.
43. Zhang, X., Zhou, L. & Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.* **19**, 3509-3519.
44. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389-3402.
45. Shuman, S. (1994). Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J. Biol. Chem.* **269**, 32678-32684.
46. Gilbert, J. M., Fowler, A., Bleibaum, J. & Clarke, S. (1988). Purification of homologous protein carboxyl methyltransferase isozymes from human and bovine erythrocytes. *Biochemistry,* **27**, 5227-5233.
47. Boggon, T. J. & Shapiro, L. (2000). Screening for phasing atoms in protein crystallography. *Structure,* **8**, 143-149.
48. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **27**, 307-326.
49. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D,* **55**, 849-861.
50. Uson, I. & Sheldrick, G. M. (1999). Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9**, 643-648.
51. Cowtan, K. D. & Main, P. (1996). Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallog. sect. D,* **52**, 43-48.
52. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.
53. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high resolution refinement. *Methods Enzymol.* **277**, 319-343.
54. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049-1067.
55. Colovos, C. & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511-1519.
56. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.
57. Lowenson, J. D. & Clarke, S. (1991). Structural elements affecting the recognition of L-isoaspartyl residues by the L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. *J. Biol. Chem.* **266**, 19396-19406.
58. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallog. sect. D,* **55**, 484-491.
59. Cools, M. & De Clercq, E. (1990). Influence of *S*-adenosylhomocysteine hydrolase inhibitors on *S*-adenosylhomocysteine and *S*-adenosylmethionine pool levels in L929 cells. *Biochem. Pharmacol.* **40**, 2259-2264.
60. Cohen, G. H. (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallog.* **30**, 1160-1161.
61. Barton, G. J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37-40.
62. Evans, S. V. (1993). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134-138.

Edited by I. A. Wilson

(Received 26 June 2001; received in revised form 14 September 2001; accepted 14 September 2001)