Sedlin) to prevent premature membrane constriction, thus allowing growth of nascent carriers and the incorporation of large PC prefibrils.

Our finding that Sedlin binds Sar1 adds a further layer to interactions between TRAPP and COPII, because Bet3 binds Sec23 (*19*), but only after Sar1 has been released from membranes and vesicles have detached from ER membranes (*20*). The Sedlin-Sar1 interaction precedes release of COPII vesicles and accelerates dissociation of Sar1 from membranes, thus allowing the next layer of interaction between COPII and TRAPP. TRAPP would then play a role in subsequent events such as Rab1 activation (*7*), thereby exhibiting the unique property of coupling the cycles of Sar1 and Rab1, two GTPases acting in series in ER-to-Golgi transport.

References and Notes

1. G. Zanetti, K. B. Pahuja, S. Studer, S. Shim, R. Schekman, *Nat. Cell Biol.* **14**, 20 (2012).
2. K. Saito *et al.*, *Cell* **136**, 891 (2009).
3. V. Malhotra, P. Erlmann, *EMBO J.* **30**, 3475 (2011).
4. D. G. Wilson *et al.*, *J. Cell Biol.* **193**, 935 (2011).
5. A. K. Gedeon *et al.*, *Nat. Genet.* **22**, 400 (1999).
6. G. E. Tiller *et al.*, *Am. J. Hum. Genet.* **68**, 1398 (2001).
7. J. Barrowman, D. Bhandari, K. Reinisch, S. Ferro-Novick, *Nat. Rev. Mol. Cell Biol.* **11**, 759 (2010).
8. J. F. Bleasel *et al.*, *J. Rheumatol.* **22**, 255 (1995).
9. A. Zankl *et al.*, *Am. J. Med. Genet. A* **129A**, 144 (2004).
10. See supplementary materials on *Science* Online.
11. A. A. Mironov *et al.*, *J. Cell Biol.* **155**, 1225 (2001).
12. P. J. Scrivens *et al.*, *Traffic* **10**, 724 (2009).
13. R. Pepperkok, M. Lowe, B. Burke, T. E. Kreis, *J. Cell Sci.* **111**, 1877 (1998).
14. K. R. Long *et al.*, *J. Cell Biol.* **190**, 115 (2010).
15. K. Bacia *et al.*, *Sci. Rep.* **1**, 17 (2011).
16. X. Bi, R. A. Corpina, J. Goldberg, *Nature* **419**, 271 (2002).
17. S. B. Jang *et al.*, *J. Biol. Chem.* **277**, 49863 (2002).
18. O. Schlenker, A. Hendricks, I. Sinning, K. Wild, *J. Biol. Chem.* **281**, 8898 (2006).
19. H. Cai *et al.*, *Nature* **445**, 941 (2007).
20. C. Lord *et al.*, *Nature* **473**, 181 (2011).

Acknowledgments: We thank A. Luini, G. D'Angelo, B. Franco, and G. Manco for insightful discussions; the Telethon EM Core Facility (project no. GTF08001) for the EM analysis; J. Gecz for the Sedlin cDNAs; R. Schekman for the Sar1A cDNA; W. Balch for the Sar1B and Sec23 cDNAs; and J. Kimura for the use of Rx chondrocytes. Supported by Telethon grants GSP08002 and GGP06166, AIRC grant IG 8623, and Programma Operativo Nazionale grant 01_00862 (M.A.D.M.), Telethon grant GGP07075 (C.W.), NIH grant NIH/AR053696 (B.M.V.), and grants from Plan Nacional (BFU2008-00414), Consolider (CSD2009-00016), Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) Grups de Recerca Emergents (SGR2009-1488; AGAUR–Catalan Government), the European Research Council (268692). (V.M.), and an AIRC fellowship (R.V.).

Supplementary Materials

www.sciencemag.org/cgi/content/full/337/6102/1668/DC1
Materials and Methods
Figs. S1 to S13
Table S1
References (*21*–*38*)

21 May 2012; accepted 17 August 2012
10.1126/science.1224947

Radical SAM-Dependent Carbon Insertion into the Nitrogenase M-Cluster

Jared A. Wiig,* Yilin Hu,*† Chi Chung Lee, Markus W. Ribbe†

The active site of nitrogenase, the M-cluster, is a metal-sulfur cluster containing a carbide at its core. Using radiolabeling experiments, we show that this carbide originates from the methyl group of *S*-adenosylmethionine (SAM) and that it is inserted into the M-cluster by the assembly protein NifB. Our SAM cleavage and deuterium substitution analyses suggest a similarity between the mechanism of carbon insertion by NifB and the proposed mechanism of RNA methylation by the radical SAM enzymes RlmN and Cfr, which involves methyl transfer from one SAM equivalent, followed by hydrogen atom abstraction from the methyl group by a 5′-deoxyadenosyl radical generated from a second SAM equivalent. This work is an initial step toward unraveling the importance of the interstitial carbide and providing insights into the nitrogenase mechanism.

The M-cluster serves as the active site of Mo-nitrogenase, where inert nitrogen is converted to bioavailable ammonia (*1*). Arguably one of the most complex protein-bound metalloclusters, the M-cluster has a core composition of 1Mo:7Fe:9S and can be viewed as $[MoFe_3S_3]$ and $[Fe_4S_3]$ subclusters bridged by three sulfide atoms (Fig. 1). In addition, an endogenous homocitrate moiety is attached to the Mo and a μ_6-light atom is coordinated at the center of the metal-sulfur core (Fig. 1). The initial discovery of the interstitial atom was exciting because of its potential relevance to the nitrogenase mechanism (*2*). Recent studies identified this light atom as carbide (*3*, *4*), prompting questions as to where this carbide originates and how it is inserted into the M-cluster.

The M-cluster is synthesized sequentially on two assembly proteins, NifB and NifEN, before it is delivered to NifDK, the catalytic component of nitrogenase (*5*). Functional analyses of NifB provided initial insights into the carbon insertion process. NifB has a CXXXCXXC signature motif that is characteristic of a family of radical *S*-adenosylmethionine (SAM or AdoMet) enzymes (*6*, *7*). Additionally, it contains sufficient ligands to potentially accommodate the entire complement of iron atoms in the M-cluster (*5*). These observations led to the proposal that NifB uses a SAM-dependent mechanism to generate the Fe-S core of the M-cluster, which is then processed on NifEN into a mature M-cluster and delivered to its target location in NifDK (*5*). Acquiring proof for this hypothesis has proven to be challenging. Initial attempts to characterize NifB from *Azotobacter vinelandii* were hampered by the instability of this protein in aqueous solutions. Recently, mimicking the natural *nifN*(3′)-(5′)*nifB* fusion in *Clostridium pasteurianum,* this problem was circumvented by fusing the *nifB* and *nifN* genes of *A. vinelandii*, resulting in the expression of a NifEN-B fusion protein consisting of two components that act sequentially in M-cluster biosynthesis (*8*).

Subsequent metal, activity, ultraviolet/visible (UV/vis), and electron paramagnetic resonance (EPR) analyses established the presence of two transient clusters, designated L- and K-cluster, respectively, on the NifEN-B fusion protein (*8*, *9*). The K-cluster consists of a pair of $[Fe_4S_4]$ clusters, whereas the L-cluster is a $[Fe_8S_9]$ cluster that closely resembles the Fe-S core of the M-cluster (*8*, *10*–*12*) (Fig. 1). Located near the $[Fe_4S_4]$ cluster at the SAM domain of NifB, the two 4Fe units of K-cluster can be coupled into an 8Fe L-cluster in the presence of SAM (*8*) (Fig. 1). Subsequently, the L cluster is transferred to and matured on NifEN upon insertion of Mo and homocitrate by NifH, an adenosine triphosphate (ATP)–dependent reductase that also serves as the obligate electron donor for NifDK during catalysis (*8*, *13*–*15*) (Fig. 1). The resultant M-cluster is then transferred from NifEN to NifDK through direct protein-protein interactions (*8*, *13*, *14*) (Fig. 1). The sequential conversion of K- to L- to M-cluster suggests that the K- and L-clusters represent consecutive snapshots of M-cluster biosynthesis. Moreover, the SAM dependence of the conversion of K- to L-cluster points to a synthetic route to bridged metalloclusters that relies on radical chemistry at the SAM domain of NifB. Given the ability of SAM to serve as a carbon donor (*16*–*19*), it is possible that the interstitial carbide is derived from SAM and that it is incorporated during the formation of the L-cluster on NifB.

The cleavage pattern of SAM upon interaction with NifEN-B provides evidence that SAM is the source of the interstitial carbide (*20*). High-performance liquid chromatography (HPLC)

Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697–3900, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: yilinh@uci.edu (Y.H.); mribbe@uci.edu (M.W.R.)

analysis shows that SAM remains intact in the absence of NifEN-B (Fig. 2A, 1) but is cleaved into two products in the presence of NifEN-B and reductant (Fig. 2A, 2). The two products display HPLC retention times characteristic of *S*-adenosyl-homocysteine (SAH or AdoHcy) and 5′-deoxyadenosine (5′-dAH) (Fig. 2A, 3 and 4). Liquid chromatography–mass spectrometry (LC-MS) analysis further confirms this identification based on mass-to-charge (*m*/*z*) ratios and fragmentation patterns of SAH and 5′-dAH (fig. S1). The same pattern of SAM cleavage was documented for RlmN and Cfr, two radical SAM methyltransferases (*16*–*18*). One mechanism proposed for the RlmN- and Cfr-catalyzed reactions (fig. S1) involves transfer of the methyl group from one equivalent of SAM to a polypeptide cysteine residue by an SN2 mechanism, followed by generation of a 5′-deoxyadenosyl radical (5′-dA•) by the reductive cleavage of a second equivalent of SAM. Subsequently, 5′-dA• abstracts a hydrogen atom from the methylcysteine, generating 5′-dAH and a presumed methylene radical on cysteine that initiates the radical-based methylation of RNA (*18*, *19*, *21*).

Deuterium substitution experiments illustrate that the reaction catalyzed by NifEN-B also involves hydrogen atom abstraction from the methyl group of SAM. As observed previously for RlmN and Cfr (*10*), incubation of NifEN-B with [methyl-d_3]-SAM, where the three hydrogen atoms of the methyl group are labeled with deuterium, results in a mixture of deuterium-enriched 5′-dAD and unlabeled 5′-dAH (Fig. 2B, right). The formation of 5′-dAD reflects the abstraction of deuterium from the labeled methyl group by 5′-dA•, whereas the detection of 5′-dAH suggests the occurrence of (i) abortive cleavage of SAM, in which the 5′-dA• abstracts a solvent-derived hydrogen atom (*17*, *18*); (ii) hydrogen atom abstraction by 5′-dA• from an unlabeled methyl group, which is generated on NifEN-B through the action of unlabeled SAM during cell growth and remains associated with NifEN-B after purification; or (iii) a combination of both events. These experiments firmly establish the initial hydrogen removal from the methyl group as a net transfer of a hydrogen atom. Continued removal of hydrogen atoms from this group, which is required for the generation of a carbide, likely occurs through proton abstraction via acid/base chemistry (*22*, *23*); alternatively, it could involve the transfer of a hydride to Fe^{3+} and the subsequent removal of protons by the metal cluster in NifEN-B, a mechanism analogous to the conversion of H_2 to H^+ by NiFe hydrogenase (*24*) and a single Fe model complex (*25*).

The similarity between NifB- and RlmN/Cfr-catalyzed reactions in using a SAM-derived methyl group as a carbon source raises the possibility of directly visualizing the entry of carbon into NifEN-B and the flow of carbon from NifEN-B onward through carbon-14 (^{14}C) labeling experiments. Such experiments can be carried out by incubating [methyl-^{14}C]-SAM with (i) NifEN-B alone or (ii) NifEN-B, NifH, and apo NifDK in a maturation assay that also contains dithionite, ATP, molybdate, and homocitrate (*13*). The former allows the SAM-dependent conversion of K-cluster to L-cluster on NifEN-B, which can be used to assess whether carbon is incorporated in the L-cluster; whereas the latter allows the assembly process to proceed with the maturation of L-cluster to M-cluster by NifH and the transfer of M-cluster to apo NifDK, which can be used to determine whether carbon is transferred together with the cluster species between these proteins (see Fig. 1). Alone, the His-tagged NifEN-B can be captured on immobilized metal affinity chromatography (IMAC) resin after incubation with excess [methyl-^{14}C]-SAM; in a maturation assay, it can be strategically combined with His- or nontagged NifH and apo NifDK, followed by the capture of His-tagged and nontagged proteins on affinity (IMAC) and anion exchange (DEAE) resin, respectively (Fig. 3A and fig. S3).

Indeed, when His-tagged NifEN-B is captured on IMAC resin after incubation with [methyl-^{14}C]-SAM, it retains radioactivity after extensive wash with buffer, suggesting a transfer of ^{14}C label from [methyl-^{14}C]-SAM to NifEN-B (Fig. 3A, 1).

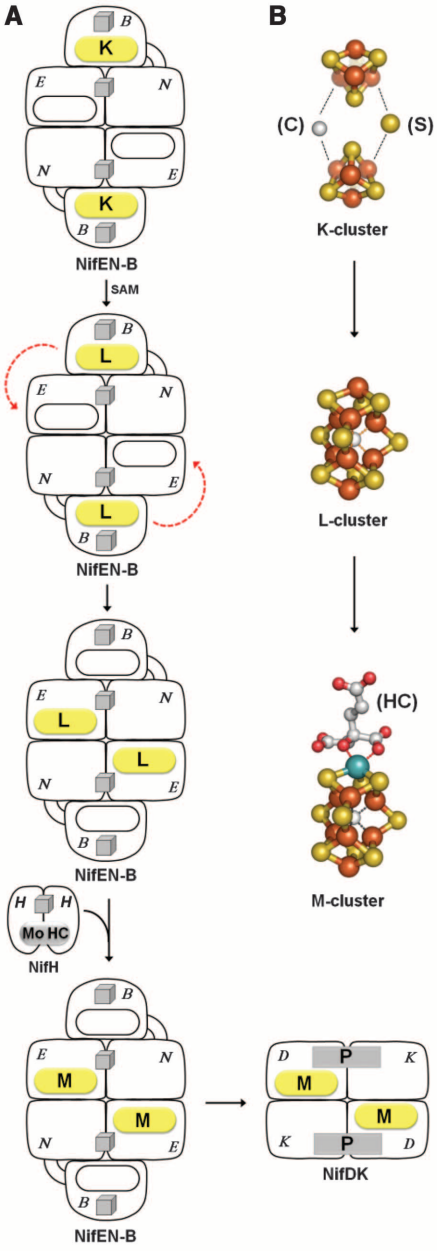

Fig. 1. (**A**) Biosynthesis of M-cluster on NifEN-B fusion protein, which involves SAM-dependent conversion of K-cluster to L-cluster on NifB, transfer of L-cluster to NifEN, maturation of L-cluster upon NifH-mediated insertion of Mo and homocitrate, and transfer of the resultant M-cluster to NifDK. (**B**) Coupling of the 4Fe units of K-cluster into an 8Fe L-cluster by carbon and sulfur insertion, and conversion of the L-cluster to a mature M-cluster upon insertion of Mo and HC (homocitrate). The L-cluster represents an all-iron homolog of the M-cluster.

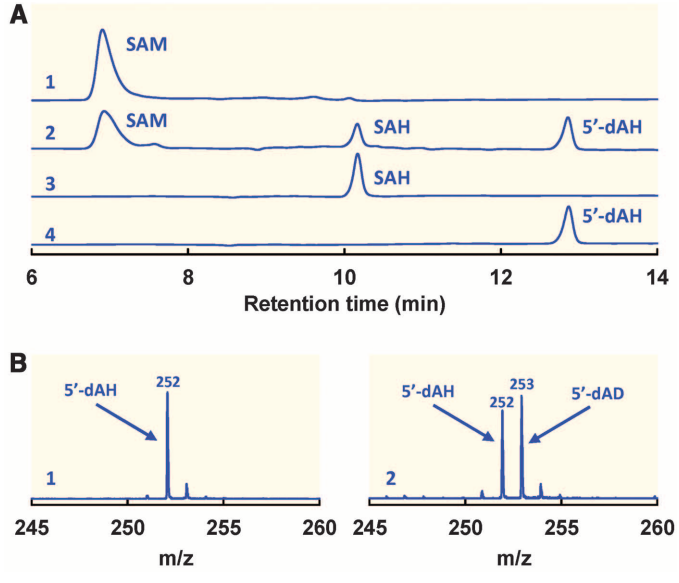

Fig. 2. (**A**) HPLC elution profiles of SAM in the (1) absence and (2) presence of NifEN-B and dithionite, compared with those of (3) SAH and (4) 5′-dAH standards. (**B**) LC-MS analysis of 5′-dAH recovered after incubation with unlabeled SAM (left) or [methyl-d_3] SAM (right). The ratio of 5′-dAH to SAH produced in the presence of NifEN-B was 1.8 ± 0.1 (N = 5 repetitions). In a maturation assay, the yields of SAH and M cluster were 2.0 ± 0.2 and 1.7 ± 0.1 μM, respectively, in the presence of 10 μM SAM and 10 μM NifEN-B (N = 5 repetitions).

After the incubation of His-tagged NifEN-B with nontagged NifH and His-tagged NifDK in a maturation assay, the ^{14}C label can be detected in the IMAC fraction (i.e., NifEN-B and NifDK) but not in the DEAE fraction (i.e., NifH), indicating that the ^{14}C label is not harbored on NifH (Fig. 3A, 2). When His-tagged NifDK is replaced by nontagged NifDK in the same maturation assay, however, the ^{14}C label appears in the DEAE fraction (i.e., NifH and NifDK), suggesting a transfer of ^{14}C label from NifEN-B to NifDK (Fig. 3A, 3). A further substitution of His-tagged NifH for nontagged NifH in this assay clearly shows the presence of radioactivity in the DEAE fraction (i.e., NifDK), providing definitive proof for the transfer of ^{14}C label to NifDK upon cluster maturation (Fig. 3A, 4). When [methyl-^{14}C]-SAM is replaced by [carboxyl-^{14}C]-SAM in the same set of experiments, radioactivity is not detected in any of the protein fractions (Fig. 3B). Thus, the radioactive labeling of NifEN-B and NifDK does not originate from the carboxyl group or from the nonspecific binding of SAM to these proteins; rather, it originates from the ^{14}C-labeled methyl group of SAM.

The ^{14}C label can be further traced to the cluster species extracted from NifEN-B and NifDK. The amount of extractable ^{14}C-labeled cluster species on NifEN-B (Fig. 3C, 1, left) is substantially decreased after the maturation of L-cluster and transfer of M-cluster from NifEN-B to apo NifDK (Fig. 3C, 2, left), and such a decrease is accompanied by an appreciable accumulation of extractable ^{14}C-labeled cluster species on NifDK after the incorporation of M-cluster (Fig. 3C, 2, middle). Activity analyses confirm that the clusters extracted from NifEN-B and NifDK are L- and M-cluster, respectively, because the former needs to be matured further on NifEN before it can be inserted into apo NifDK (Fig. 3D, 4 and 5), whereas the latter can be used directly to reconstitute apo NifDK (Fig. 3D, 6). SDS–polyacrylamide gel electrophoresis analysis further demonstrates the strict association of ^{14}C label with the cluster species, because no radioactivity is detected in the polypeptides of [methyl-^{14}C]-SAM-treated NifEN-B or NifDK (fig. S4). The absence of ^{14}C label from NifEN-B polypeptides also implies a direct transfer of methyl-derived carbon species to the K-cluster. This is supported by data from a posttranslational modification (PTM) analysis, which reveal the absence of methyl or methylene group from NifEN-B polypeptides after treatment with SAM (*26*).

Based on these observations, we propose two plausible mechanisms for the carbon insertion by NifB, which are clearly distinct in the requirement for reductant in the initial step of methyl transfer (Fig. 4). The first mechanism (Fig. 4A) begins with the transfer of a methyl group from SAM to a sulfide of the K-cluster through a reductant-independent SN2 mechanism, generating SAH and a cluster that at least transiently

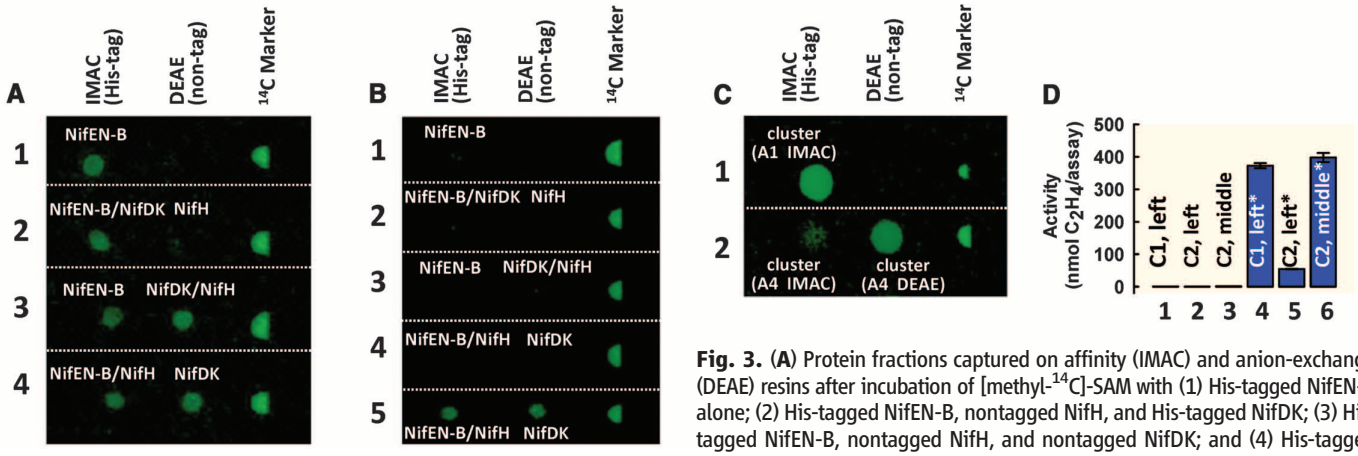

Fig. 3. (**A**) Protein fractions captured on affinity (IMAC) and anion-exchange (DEAE) resins after incubation of [methyl-^{14}C]-SAM with (1) His-tagged NifEN-B alone; (2) His-tagged NifEN-B, nontagged NifH, and His-tagged NifDK; (3) His-tagged NifEN-B, nontagged NifH, and nontagged NifDK; and (4) His-tagged NifEN-B, His-tagged NifH, and nontagged NifDK. Assays 2 to 4 also contained dithionite, ATP, molybdate, and homocitrate. (**B**) Experiments (B, 1 to 4) were identical to those in (A, 1 to 4), except that [carboxyl-^{14}C]-SAM was used instead of [methyl-^{14}C]-SAM. Experiment (A, 4) was included as a positive control here (B, 5). (**C**) Clusters extracted from (1) the IMAC fraction in (A, 1); and (2) the IMAC fraction (left) and DEAE (middle) fraction in (A, 4). (**D**) Activities of C_2H_2 reduction by extracted clusters alone (1 to 3), after maturation and transfer of clusters to apo NifDK (4 and 5), or upon direct transfer of cluster to apo NifDK (6). Clusters used for activity analyses were from C1, left (1 and 4); C2, left (2 and 5); and C2, middle (3 and 6), respectively. Activities of clusters upon maturation and/or incorporation into apo NifDK (4-6) are indicated by *.

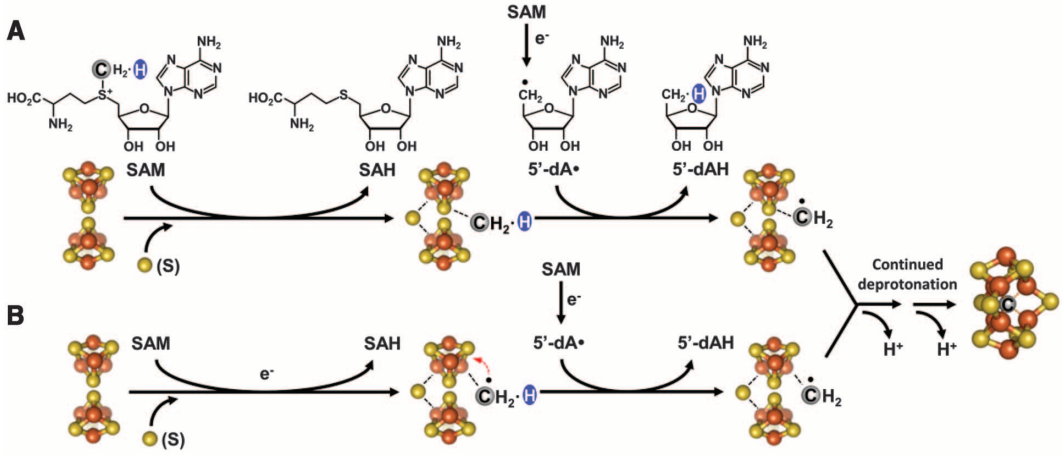

Fig. 4. Proposed mechanisms of carbide insertion involve (**A**) transfer of a methyl group from SAM to the 4Fe K-cluster through an SN2 mechanism, followed by reductive cleavage of a second equivalent of SAM to generate 5-dA•, which abstracts a hydrogen atom from the methyl group on the K-cluster; or (**B**) reductive cleavage of a methyl group from SAM, followed by transfer of the resultant methyl radical to the 4Fe K-cluster and abstraction of a hydrogen atom of this group by 5-dA•. Continued deprotonation of the C intermediate in (A) or (B) must occur concomitant with its incorporation into the 8Fe L-cluster as an interstitial carbide atom. Atoms of the clusters are colored as follows: Fe, orange; S, yellow; C, gray.

contains a methanethiol ligand. Then, reductive cleavage of a second equivalent of SAM gives rise to a 5′-dA•, which abstracts a hydrogen atom from the methyl group, generating 5′-dAH and a methylene thiol radical. This scenario is consistent with the fact that Fe^{2+} in the K-cluster is not sufficiently nucleophilic for the initial methyl group transfer to generate an Fe-C bond (*27*). Consequently, it must involve a switch of the ligation of the C intermediate to Fe atom(s), as well as the removal of the other two hydrogen atoms of this intermediate, which eventually leads to the formation of an interstitial carbide atom that is coordinated by six Fe atoms. The second mechanism (Fig. 4B) starts with the reductive cleavage of a methyl group from SAM, which generates a transient methyl radical that could be captured on Fe^{2+} of the K-cluster to directly form an Fe-C bond. This scenario is analogous to the capture of C radicals by Co^{2+} in vitamin B12 radical chemistry (*28*, *29*), and it eliminates the need for ligand exchange of the C intermediate. The subsequent hydrogen transfer could proceed in a manner identical to that in the first mechanism, which involves the initial hydrogen atom abstraction by 5′-dA• and the subsequent proton abstraction that leads to the formation of carbide.

It should be noted that the radical SAM-based carbide insertion is an integral part of a synthetic strategy to couple the two 4Fe units of K-cluster into an 8Fe L-cluster (Fig. 4). It is accompanied by the incorporation of a cluster-bridging sulfide, as well as the structural rearrangements of Fe-S clusters through the facile ligand replacement of their tetrahedral Fe^{2+}/Fe^{3+} atoms by an addition/elimination mechanism (*30*, *31*). Details of this process, such as whether the initial methyl transfer requires a reductant and how the starting methyl group is processed into a carbide atom, are yet to be defined. Nevertheless, this work is a step toward unraveling the structure-function relationship of the interstitial carbide, which could provide insights into the mechanism of nitrogenase.

References and Notes

1. B. K. Burgess, D. J. Lowe, *Chem. Rev.* **96**, 2983 (1996).
2. O. Einsle *et al.*, *Science* **297**, 1696 (2002).
3. K. M. Lancaster *et al.*, *Science* **334**, 974 (2011).
4. T. Spatzal *et al.*, *Science* **334**, 940 (2011).
5. G. Schwarz, R. R. Mendel, M. W. Ribbe, *Nature* **460**, 839 (2009).
6. H. J. Sofia, G. Chen, B. G. Hetzler, J. F. Reyes-Spindola, N. E. Miller, *Nucleic Acids Res.* **29**, 1097 (2001).
7. P. A. Frey, A. D. Hegeman, F. J. Ruzicka, *Crit. Rev. Biochem. Mol. Biol.* **43**, 63 (2008).
8. J. A. Wiig, Y. Hu, M. W. Ribbe, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8623 (2011).
9. The reduced K-cluster shows an $S = 1/2$ signal that is characteristic of $[Fe_4S_4]^+$ clusters, whereas the oxidized L-cluster shows a $g = 1.92$ signal that is unique to the $[Fe_8S_9]$ precursor of M-cluster (*8*). *S* is spin quantum number; *g* is spectroscopic splitting factor. The assignment of K-cluster as a pair of $[Fe_4S_4]$ clusters was based on metal determination and quantitative EPR and UV/vis analyses (*8*), whereas the determination of L-cluster as a $[Fe_8S_9]$ homolog to the M-cluster was based on x-ray absorption spectroscopy, extended x-ray absorption fine structure, and x-ray crystallographic studies (*10*–*12*). The conversion of K- to L- to M-cluster on NifEN-B was monitored by activity and EPR analyses (*8*).
10. M. C. Corbett *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1238 (2006).
11. A. W. Fay *et al.*, *Angew. Chem. Int. Ed. Engl.* **50**, 7787 (2011).
12. J. T. Kaiser, Y. Hu, J. A. Wiig, D. C. Rees, M. W. Ribbe, *Science* **331**, 91 (2011).
13. Y. Hu, A. W. Fay, M. W. Ribbe, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3236 (2005).
14. Y. Hu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17119 (2006).
15. Y. Hu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17125 (2006).
16. F. Yan *et al.*, *J. Am. Chem. Soc.* **132**, 3953 (2010).
17. F. Yan, D. G. Fujimori, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3930 (2011).
18. T. L. Grove *et al.*, *Science* **332**, 604 (2011).
19. A. K. Boal *et al.*, *Science* **332**, 1089 (2011).
20. Materials and methods are available as supplementary materials on *Science* Online.
21. An alternative mechanism was proposed for RNA methylation by RlmN, which involves hydrogen atom abstraction from the methyl group of SAM by 5′-dA•, followed by transfer of the resultant methylene radical to an amidine carbon of the substrate (*16*).
22. M. W. van der Kamp, J. Żurek, F. R. Manby, J. N. Harvey, A. J. Mulholland, *J. Phys. Chem. B* **114**, 11303 (2010).
23. F. C. Hartman, E. H. Lee, *J. Biol. Chem.* **246**, 11784 (1989).
24. M. Pavlov, P. E. M. Siegbahn, M. R. A. Blomberg, R. H. Crabtree, *J. Am. Chem. Soc.* **120**, 548 (1998).
25. X. Yang, M. B. Hall, *J. Am. Chem. Soc.* **131**, 10901 (2009).
26. PTM analyses were performed as described previously (*18*) at the Biomolecular Mass Spectrometry Facility of University of California–San Diego and the Mass Spectrometry Facility of University of California–Irvine.
27. R. D. Hancock, A. E. Martell, *Chem. Rev.* **89**, 1875 (1989).
28. G. H. Reed, *Curr. Opin. Chem. Biol.* **8**, 477 (2004).
29. R. Banerjee, *Chem. Rev.* **103**, 2083 (2003).
30. G. R. Dukes, R. H. Holm, *J. Am. Chem. Soc.* **97**, 528 (1975).
31. R. C. Job, T. C. Bruice, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2478 (1975).

Acknowledgments: This work was supported by National Institutes of Health grant GM 67626 (M.W.R.).

Supplementary Materials

www.sciencemag.org/cgi/content/full/337/6102/1672/DC1
Materials and Methods
Figs. S1 to S4
References (*32*–*36*)

11 May 2012; accepted 27 August 2012
10.1126/science.1224603

Evidence of Abundant Purifying Selection in Humans for Recently Acquired Regulatory Functions

Lucas D. Ward[1,2] and Manolis Kellis[1,2]*

Although only 5% of the human genome is conserved across mammals, a substantially larger portion is biochemically active, raising the question of whether the additional elements evolve neutrally or confer a lineage-specific fitness advantage. To address this question, we integrate human variation information from the 1000 Genomes Project and activity data from the ENCODE Project. A broad range of transcribed and regulatory nonconserved elements show decreased human diversity, suggesting lineage-specific purifying selection. Conversely, conserved elements lacking activity show increased human diversity, suggesting that some recently became nonfunctional. Regulatory elements under human constraint in nonconserved regions were found near color vision and nerve-growth genes, consistent with purifying selection for recently evolved functions. Our results suggest continued turnover in regulatory regions, with at least an additional 4% of the human genome subject to lineage-specific constraint.

Initial sequencing of the human genome revealed that 98.5% of human DNA does not code for protein (*1*), raising the question of what fraction of the remaining genome is functional. Mammalian conservation suggests that ~5% of the human genome (*2*, *3*) is conserved due to noncoding and regulatory roles, but more than 80% is transcribed, bound by a regulator, or associated with chromatin states suggestive of regulatory functions (*4*–*6*). This discrepancy may result from nonconsequential biochemical activity or lineage-specific constraint (*7*, *8*). Similarly, evolutionary turnover in regulatory regions (*9*–*11*) may be due to nonconsequential activity in neutrally evolving regions in each species or turnover in functional elements associated with turnover in activity. To resolve these questions, we need new methods for measuring constraint within a species, rather than between species.

Single-nucleotide polymorphisms (SNPs) within human populations have been identified only every 153 bases on average (*12*), compared with 4.5 substitutions per site among the genomes of 29 mammals (*2*), making it impossible to detect individual constrained elements (*13*). Instead, aggregate measures of human diversity across thousands of dispersed elements are needed. Such

[1]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]The Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: manoli@mit.edu