Functional replacement of histidine in proteins to generate noncanonical amino acid dependent organisms

Fei Gan, Renhe Liu, Feng Wang, and Peter G Schultz

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.7b13452 • Publication Date (Web): 06 Mar 2018

Functional replacement of histidine in proteins to generate noncanonical amino acid dependent organisms

[1] Fei Gan, [1] Renhe Liu, [1] Feng Wang, and [1, 2]Peter G. Schultz*

[1]California Institute for Biomedical Research (Calibr), La Jolla, California 92037, United States

[2]Department of Chemistry and the Skaggs Institute for Chemical Biology, the Scripps Research Institute, La Jolla, California 92037, United States.

Supporting Information Placeholder

ABSTRACT: Simple strategies to produce organisms whose growth is strictly dependent on the presence of a noncanonical amino acid are useful for the generation of live vaccines and the biological containment of recombinant organisms. To this end, we report an approach based on genetically replacing key histidine (His) residues in essential proteins with functional His analogs. We demonstrate that 3-methyl-L-histidine (MeH) functionally substitutes for a key metal binding ligand, H264, in the zinc-containing metalloenzyme mannose-6-phosphate isomerase (ManA). An evolved variant, Opt5, harboring both N262S and H264MeH substitutions exhibited comparable activities to wild type ManA. An engineered *Escherichia coli* strain containing the ManA variant Opt5 was strictly dependent on MeH for growth with an extremely low reversion rate. This straightforward strategy should be applicable to other metallo- or non-metalloproteins that contain essential His residues.

Genetically encoded noncanonical amino acids (ncAAs) are a powerful tool for probing the structure and function of proteins, as well as creating proteins with novel or advanced activities[1-3]. Recently, organisms strictly depending on the presence of ncAAs have been created for biological containment and live attenuated vaccines[4-8]. Engineered organisms grow only when the specified ncAA is available to translate an in-frame nonsense codon to produce a fully functional protein. However, the incorporation of ncAAs at catalytically active sites or at key structural elements usually adversely affects protein activity, whereas substitution at permissive sites suffers from high escape frequencies due to naturally occurring mutations. Therefore, additional strategies to create ncAA-dependent organisms with normal growth and low escape frequencies are needed.

Metalloproteins account for one third to one half of naturally occurring proteins, and metal ions can play key catalytic roles or act as essential structural elements[9,10]. In particular, zinc ions are required for many hydrolytic and redox enzymes, and histidine (His) is the most common metal coordinating ligand [11-14]. It is also noted that two single mutations are required to convert a nonsense codon (e.g.,UAA, UAG) to either of the native codons for His (CAU and CAC). Therefore, substitution of a metal-binding His in an essential protein with a functional His analog, encoded by an amber or ochre nonsense codon, should result in ncAA conditional cell survival with a low reversion rate, as reversion to His would be very unlikely.

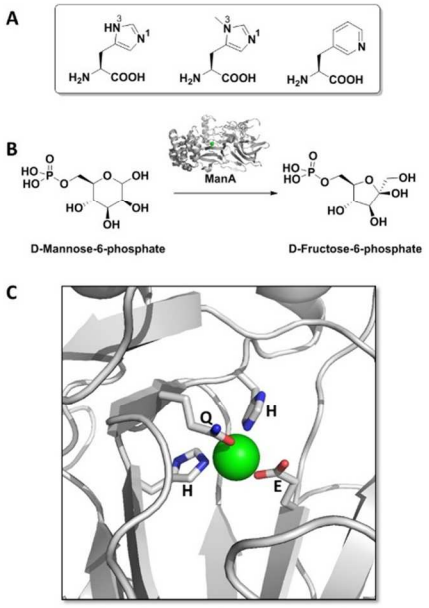


Figure 1. ManA as a model system and ncAAs used in this study. (A) Canonical amino acid His (left) and its selected analogs 3-methyl-L-Histidine (MeH, middle) and 3-(3-pyridyl)-alanine (PyA, right). (B) The reaction catalyzed by ManA. (C) The Zn^{2+}-binding pocket of ManA from *Candida albicans* (PDB: 1PMI, Zn^{2+} cofactor shown in green).

Metal-chelating ncAAs including 2,2'-bipyridin-5-yl-alanine (Bpy-Ala) and 8-hydroxyquinolin-3-yl-alanine have been genetically encoded to generate a Cu^{2+}-dependent DNA cleaving protein[15], metal-binding biophysical probes[16], novel iron(II) binding zinc finger-like motifs[17], catalytic dimer interfaces[18,19], and stabilized trimeric helical motifs[20]. Engineering of Bpy-Ala dependent metalloproteins with atom level accuracy was accomplished by Miller and coworkers using computational design method[21,22]. However, it remains challenging to introduce these bulky bidentate ncAAs at enzyme active sites and retain native activity[17,21,22]. Structurally similar analogs to His, such as 3-methyl-L-histidine and 3-(3-pyridyl)-L-alanine (Figure 1A), are potential functional replacements for His with minimal perturbation on protein structure and activity[23-25]. To demonstrate this notion, we selected the metalloenzyme mannose-6-phosphate isomerase (ManA) as a model. ManA is an essential enzyme in various pathogens, such as *Candida albicans* and *Mycobacterium tuberculosis* (*Mtb*), and catalyzes the interconversion of D-mannose-6-phosphate (M6P) and D-fructose-6-phosphate (F6P) (Figure 1B) which are involved in multiple metabolic pathways, cell wall formation, and capsular polysaccharide biosynthesis[26,27]. The activity of ManA relies on a metal cofactor Zn^{2+} [28]. Two His (H), one Glu (E), and one Gln (Q) (or Lys in certain organisms) commonly serve as Zn^{2+}-binding ligands and are highly conserved in ManA homologs (Figure 1C and Figure S1). In particular *Mtb* ManA contains the two His residues (H99 and H264, Figure S1) and H99 has been proposed to act as a general acid in the ring opening of substrate[28,29].

To incorporate His analogs in response to an amber codon at site H99 or H264 of *Mtb* ManA, we first constructed a recombinant plasmid pWT for expression of wild type (WT) ManA under the T5 promoter. The codons at site 99 and 264 were then independently mutated to TAG to afford plasmids pH99TAG and pH264TAG for expression of ManA variants H99X and H264X (X: MeH or PyA), respectively. Each of the constructs was individually co-transformed into *Escherichia coli (E. coli)* DH10B cells with plasmid pUltra-HRS, which harbors the genes encoding the cognate aaRS and tRNA for specific incorporation of the His analogues[30]. The aminoacyl-tRNA synthase HRS is polyspecific and uses both MeH and PyA as substrates, but does not recognize His. Protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) in the presence of 1 mM MeH or PyA. The yields for ncAA-containing variants after purification were ~ 10-30 μg/L, while a higher yield was obtained for WT ManA (250 μg/L). Mass spectrometry analysis showed that the observed masses for WT, H99MeH, H99PyA, H264MeH, and H264PyA were 44194 Da, 44208 Da, 44205 Da, 44208 Da, and 44205 Da respectively (Figure S2), which were consistent with the expected masses. These data indicate that both MeH and PyA can be specifically incorporated in response to an amber codon at either H99 or H264 site.

We next investigated the activities of ManA variants H99X and H264X using a growth assay with an *E.coli* mutant strain lacking endogenous ManA activity. Due to the inactivation of its native *manA* gene, *E.coli ΔmanA* mutant strain is unable to grow on D-mannose as the sole carbon source (Figure 2A). Heterologous expression of WT *Mtb* ManA in this strain restored the capacity to use D-mannose for growth (Figure 2B). To determine the activities of H99X and H264X variants, we co-transformed plasmid pH99TAG or pH264TAG individually with pUltra-HRS into *E.coli ΔmanA* cells, and grew the resulting strain (Ec.ΔmanA.H99X or Ec.ΔmanA.H264X) on D-mannose minimal medium (termed M9M medium) supplemented with 1 mM IPTG, appropriate antibiotics, and 1 mM MeH or PyA (Materials and Methods in SI). No growth was observed for Ec.ΔmanA.H99TAG cells in the presence of MeH or PyA (Figure 2C and 2E), suggesting neither ncAA can functionally substitute for H99. Strain Ec.ΔmanA.H264TAG also did not grow in the presence of PyA (Figure 2D); however, MeH can functionally replace H264 and growth was observed for this strain when supplemented with 1 mM MeH (Figure 2F). These results agree with previous studies proposing a dual role of H99 involving both imidazole nitrogens whereas the N^1 of H264 chelates Zn^{2+} [28,29].

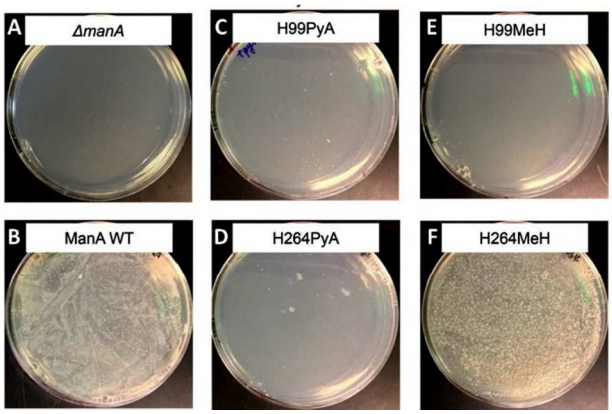


Figure 2. A ManA deficient *E.coli* strain expressing *Mtb* WT ManA and ncAA-containing variants. (A) and (B) Growth of the *E.coli ΔmanA* mutant and strain expressing *Mtb* wild type ManA (ManA WT) respectively. (C) and (E) Growth of Ec.ΔmanA.H99TAG strain in the presence of 1 mM PyA (H99PyA) and 1 mM MeH (H99MeH) on M9M selection medium respectively. (D) and (F) Growth of Ec.ΔmanA.H264TAG strain in the presence of 1 mM PyA (H264PyA) and 1 mM MeH (H264MeH) on M9M selection medium respectively.

Considering several canonical amino acid residues (e.g., Glu, Cys) can function as metal binding ligands, we next determined if H264 can be replaced by any of the common amino acids. We therefore mutated H264 to each of the other 19 canonical amino acids. None of the canonical ManA variants restored growth of the *E.coli ΔmanA* mutant on D-mannose (Figure S3). These data demonstrate that any canonical substitution at 264 site, which could occur through a single mutation of an amber codon, leads to inactivated

ManA. A low reversion rate is thus expected for the engineered strain containing the His264MeH ManA variant. To test this notion, we grew Ec.ΔmanA.H264MeH cells continuously in the presence of 1 mM MeH, and screened $1x10^{11}$ cells for survival on non-permissive medium (in the absence of MeH) for a 2-week incubation at 30°C. The reversion rate determined for Ec.ΔmanA.H264MeH was less than 10^{-11} escapees per colony forming unit. Importantly, these values are below the acceptable reversion frequency (10^{-8}) for biocontainment requirement in the guidelines of the National Institutes of Health (NIH)[31].

Despite growing continuously in the presence of MeH, the Ec.ΔmanA.H264MeH strain grew slower than the Ec.ΔmanA.WT strain (Figure S4). Western blot analysis showed a lower amount of variant H264MeH relative to WT ManA was expressed in cells (Figure S5) possibly due to the efficiency of MeH incorporation by the cognate HRS/tRNA pair at this site. Considering ManA plays critical roles in cellular metabolism, a lower ManA content could lead to a decrease in growth rate. Another possibility is that the methyl group at N^3 of MeH might slightly alter the geometry of the Zn^{2+} site, leading to reduced activity. We therefore performed random mutagenesis on ManA coding sequence by error prone PCR in an effort to evolve an H264MeH variant with increased activity. The amount of template pH264TAG was adjusted to achieve a high mutation rate (9-16 mutations per kb) during error prone PCR. The resulting library (pH264TAG-RM) was introduced into *E.coli ΔmanA* mutant cells harboring the pUltra-HRS plasmid, followed by screening on MeH-containing M9M selection medium. Sequencing of colonies exhibiting faster growth revealed a single mutation N262S along with the amber codon at H264 site (Figure S6A) (denoted Opt5). The observed mass of purified Opt5 protein was 44181 Da, which was consistent with expected mass of a ManA variant with both N262S and H264MeH substitutions (Figure S6B).

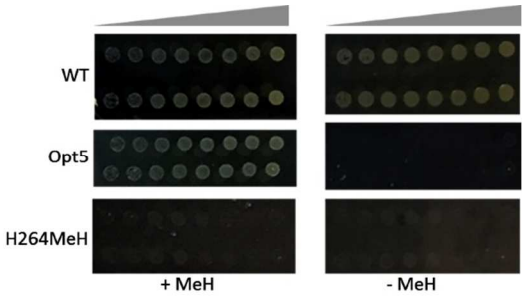


Figure 3. Growth of *E.coli ΔmanA* cells expressing WT and variants Opt5 and H264MeH on solid M9M selective medium in the presence or absence of 1 mM MeH. Two single colonies for each were assayed. Equal number of cells (after 10-fold serial dilution) were spotted and cultivated for 24 hours.

We further characterized the ManA variant Opt5 in comparison with WT ManA and variant H264MeH. The Ec.ΔmanA.Opt5 strain grew only in the presence of 1 mM MeH on selective medium with improved rate compared with Ec.ΔmanA.H264MeH strain (Figure 3 and Figure S4). Given variant Opt5 was expressed at a similar level as H264MeH (Figure S5), we reasoned that the faster cell growth could be due to an improvement in activity. The catalytic activities of purified WT, Opt5 and H264MeH ManA proteins were then measured using a coupled reaction monitoring absorbance at 340 nm due to reduction of $NADP^+$ (Figure 4, Figures S7 and S8). The k_{cat} and K_m values of variant Opt5 ($k_{cat} = 29.5 \pm 1.0\ s^{-1}$; $K_m = 1.2 \pm 0.1$ mM; $k_{cat}/K_m = 24.6\ mM^{-1}\ s^{-1}$) are comparable to those of WT enzyme ($k_{cat} = 29.3 \pm 1.2\ s^{-1}$; $K_m = 0.7 \pm 0.1$ mM; $k_{cat}/K_m = 41.9\ mM^{-1}\ s^{-1}$). We could not obtain k_{cat} and K_m values for H264MeH which exhibited a low activity. To explain the increased activity of the N262S mutation, a model of *Mtb* ManA with I-TASSER[32] was constructed (Figure S9). N262 is located in a loop that connects the β-sheet containing H264, and is in close proximity (~4 Å) to the N^3 atom of H264. It is likely that this mutation helps sterically to accommodate the H264MeH substitution, and potentially improves the geometry of the Zn^{2+} coordination site. Although one might like to further increase the growth rate of the Opt5 strain to that of the WT strain by improving protein expression levels, a conditional vaccine will be produced ex vivo and therefore the observed growth rate should be sufficient.

Perhaps the most critical experiment for a conditional vaccine is to evaluate whether the N262S mutation affects the low escape frequency. We therefore screened $1x10^{11}$ Ec.ΔmanA. Opt5 cells for survival on non-permissive medium (in the absence of MeH) for a 2-weeks incubation at 30°C. The reversion rate was less than 10^{-11} escapees per colony forming unit, comparable to that of strain Ec.ΔmanA.H264MeH.

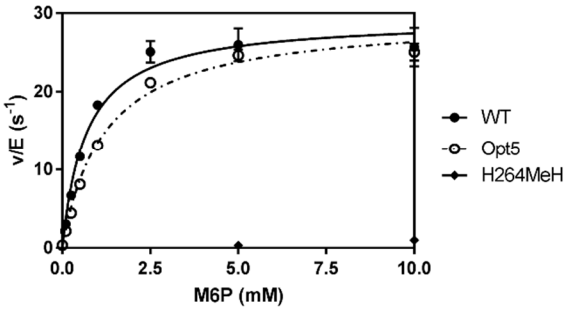


Figure 4. Catalytic activity of WT ManA and variants. Activity was measured by monitoring A_{340nm} from reduction of $NADP^+$ in a coupled reaction. Error bars represent standard deviation of triplicates. M6P: D-mannose-6-phosphate.

In summary, we describe a strategy for creating ncAA dependent organisms by genetically substituting key His residues in essential proteins with functional His analogs. This strategy uses an orthogonal tRNA/aaRS pair to insert MeH at in-frame amber codons, and thus requires minimal genetic modification of the host organisms. Notably, this strategy results in stringent ncAA dependence and extremely low reversion rates. Although we chose a His residue that coordinates a catalytic Zn^{2+} ion, this approach can likely be applied to structural Zn^{2+} sites, and His residues that play other critical roles in catalysis. We are currently exploring the feasibility of using this system to create ncAA conditional live vaccines.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supplementary Materials and Methods, Figures s1-s9, and Table S1.

AUTHOR INFORMATION

Corresponding Author

*schultz@scripps.edu.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

The authors thank Dr. Jeremy Mills for initial computional analysis. The authors also acknowledge Dr. Weimin Xuan and Dr. Minseob Koh for discussions and Kristen Williams for assistance in manuscript preparation. This work was supported, in part, by Bill and Melinda Gates Foundation.

REFERENCES

(1) Young, T. S.; Schultz, P. G. *J. Biol. Chem.* **2010**, *285*, 11039.
(2) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79*, 413.
(3) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225.
(4) Xuan, W.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 9170.
(5) Koh, M.; Nasertorabi, F.; Han, G. W.; Stevens, R. C.; Schultz, P. G. *J. Am. Chem. Soc.* **2017**, *139*, 5728.
(6) Si, L.; Xu, H.; Zhou, X.; Zhang, Z.; Tian, Z.; Wang, Y.; Wu, Y.; Zhang, B.; Niu, Z.; Zhang, C.; Fu, G.; Xiao, S.; Xia, Q.; Zhang, L.; Zhou, D. *Science* **2016**, *354*, 1170.
(7) Rovner, A. J.; Haimovich, A. D.; Katz, S. R.; Li, Z.; Grome, M. W.; Gassaway, B. M.; Amiram, M.; Patel, J. R.; Gallagher, R. R.; Rinehart, J.; Isaacs, F. J. *Nature* **2015**, *518*, 89.
(8) Mandell, D. J.; Lajoie, M. J.; Mee, M. T.; Takeuchi, R.; Kuznetsov, G.; Norville, J. E.; Gregg, C. J.; Stoddard, B. L.; Church, G. M. *Nature* **2015**, *518*, 55.
(9) Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J. *Nature* **2009**, *460*, 823.
(10) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855.
(11) Holliday, G. L.; Mitchell, J. B.; Thornton, J. M. *J. Mol. Biol.* **2009**, *390*, 560.
(12) Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M. *J. Mol. Biol.* **2002**, *324*, 105.
(13) McCall, K. A.; Huang, C.; Fierke, C. A. *J. Nutr.* **2000**, *130*, 1437s.
(14) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647.
(15) Lee, H. S.; Schultz, P. G. *J. Am. Chem. Soc*. **2008**, *130*, 13194.
(16) Lee, H. S.; Spraggon, G.; Schultz, P. G.; Wang, F. *J. Am. Chem. Soc.* **2009**, *131*, 2481.
(17) Kang, M.; Light, K.; Ai, H. W.; Shen, W.; Kim, C. H.; Chen, P. R.; Lee, H. S.; Solomon, E. I.; Schultz, P. G. *Chembiochem* **2014**, *15*, 822.
(18) Bos, J.; Browne, W. R.; Driessen, A. J. M.; Roelfes, G. *J. Am. Chem. Soc.* **2015**, *137*, 9796.
(19) Bos, J.; Fusetti, F.; Driessen, A. J.; Roelfes, G. *Angew. Chem., Int. Ed.* **2012**, *51*, 7472.
(20) Luo, X.; Wang, T. A.; Zhang, Y.; Wang, F.; Schultz, P. G. *Cell Chem. Biol.* **2016**, *23*, 1098.
(21) Mills, J. H.; Khare, S. D.; Bolduc, J. M.; Forouhar, F.; Mulligan, V. K.; Lew, S.; Seetharaman, J.; Tong, L.; Stoddard, B. L.; Baker, D. *J. Am. Chem. Soc.* **2013**, *135*, 10.1021/ja403503m.
(22) Mills, J. H.; Sheffler, W.; Ener, M. E.; Almhjell, P. J.; Oberdorfer, G.; Pereira, J. H.; Parmeggiani, F.; Sankaran, B.; Zwart, P. H.; Baker, D. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 15012.
(23) Pott, M.; Hayashi, T.; Mori, T.; Mittl, P. R. E.; Green, A. P.; Hilvert, D. *J. Am. Chem. Soc.* **2018**.
(24) Sharma, V.; Wang, Y. S.; Liu, W. R. *ACS Chem. Biol*. **2016**, *11*, 3305.
(25) Green, A. P.; Hayashi, T.; Mittl, P. R.; Hilvert, D. *J. Am. Chem. Soc.* **2016**, *138*, 11344.
(26) Griffin, J. E.; Gawronski, J. D.; Dejesus, M. A.; Ioerger, T. R.; Akerley, B. J.; Sassetti, C. M. *PLoS Pathog.* **2011**, *7*, e1002251.
(27) Cleasby, A.; Wonacott, A.; Skarzynski, T.; Hubbard, R. E.; Davies, G. J.; Proudfoot, A. E.; Bernard, A. R.; Payton, M. A.; Wells, T. N. *Nat. Struct. Biol.* **1996**, *3*, 470.
(28) Gracy, R. W.; Noltmann, E. A. *J Biol Chem* **1968**, *243*, 5410.
(29) Sagurthi, S. R.; Gowda, G.; Savithri, H. S.; Murthy, M. R. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *65*, 724.
(30) Xiao, H.; Peters, F. B.; Yang, P. Y.; Reed, S.; Chittuluru, J. R.; Schultz, P. G. *ACS Chem. Biol.* **2014**, *9*, 1092.
(31) NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules; Office of Science Policy, National Institutes of Health: Bethesda, MD, 2016.
(32) Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. *Nat. Methods* **2015**, *12*, 7.

SYNOPSIS TOC (Word Style "SN_Synopsis_TOC"). If you are submitting your paper to a journal that requires a synopsis graphic and/or synopsis paragraph, see the Instructions for Authors on the journal's homepage for a description of what needs to be provided and for the size requirements of the artwork.

To format double-column figures, schemes, charts, and tables, use the following instructions:

Place the insertion point where you want to change the number of columns
From the **Insert** menu, choose **Break**
Under **Sections**, choose **Continuous**
Make sure the insertion point is in the new section. From the **Format** menu, choose **Columns**
In the **Number of Columns** box, type **1**
Choose the **OK** button

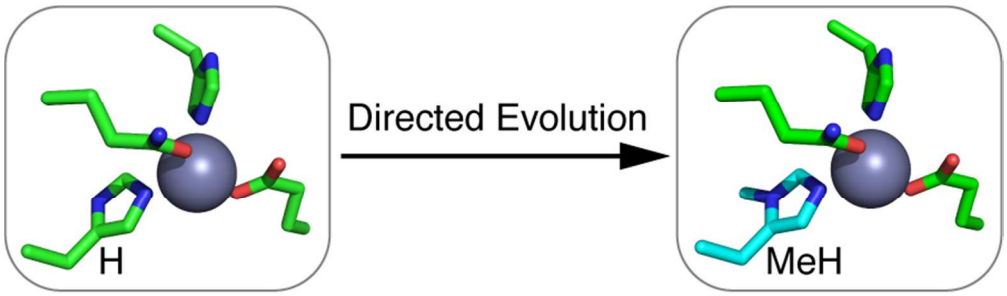
Now your page is set up so that figures, schemes, charts, and tables can span two columns. These must appear at the top of the page. Be sure to add another section break after the table and change it back to two columns with a spacing of 0.33 in.

Table 1. Example of a Double-Column Table

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8

Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

Insert Table of Contents artwork here



TOC Image

76x21mm (300 x 300 DPI)