J|A|C|S
ARTICLES
Published on Web 09/04/2004

The Mechanism of Nitrogenase. Computed Details of the Site and Geometry of Binding of Alkyne and Alkene Substrates and Intermediates

Ian Dance

Contribution from the School of Chemical Sciences, University of New South Wales, Sydney 2052, Australia

Received April 2, 2004; E-mail: I.Dance@unsw.edu.au

Abstract: The chemical mechanism by which the enzyme nitrogenase effects the remarkable reduction of N_2 to NH_3 under ambient conditions continues to be enigmatic, because no intermediate has been observed directly. Recent experimental investigation of the enzymatic consequences of the valine ⇒ alanine modification of residue α-70 of the component MoFe protein on the reduction of alkynes, together with EPR and ENDOR spectroscopic characterization of a trappable intermediate in the reduction of propargyl alcohol or propargyl amine ($HC{\equiv}C{-}CH_2OH/NH_2$), has localized the site of binding and reduction of these substrates on the FeMo-cofactor and led to proposed η^2-Fe coordination geometry. Here these experimental data are modeled using density functional calculations of the allyl alcohol/amine intermediates and the propargyl alcohol/amine reactants coordinated to the FeMo-cofactor, together with force-field calculations of the interactions of these models with the surrounding MoFe protein. The results support and elaborate the earlier proposals, with the most probable binding site and geometry being η^2-coordination at Fe6 of the FeMo-cofactor (crystal structure 1M1N in the Protein Database), in a position that is intermediate between the exo and endo coordination extremes at Fe6. The models described account for (1) the steric influence of the α-70 residue, (2) the crucial hydrogen bonding with Nε of α-195^{His}, (3) the spectroscopic symmetry of the allyl−alcohol intermediate, and (4) the preferential stabilization of the allyl alcohol/amine relative to propargyl alcohol/amine. Alternative binding sites and geometries for ethyne and ethene, relevant to the wild-type protein, are described. This model defines the location and scene for detailed investigation of the mechanism of nitrogenase.

Introduction

The nitrogenase enzymes are responsible for about half of the fixation of inert N_2 into forms used to sustain life on earth. From the chemical research perspective there is an intriguing question: what chemical mechanism allows enzymatic reduction of N_2 under mild conditions of ca. 20 °C and 0.8 atm, when the best industrial processes for nitrogen fixation use 400−500 °C and 400−1000 atm? Despite research over many years and much understanding of the enzymes involved,[1−4] we still do not have an understanding at the atomic level of how the enzyme can activate such an inert molecule under remarkably mild conditions.

Research on this question became focused in 1992, when the unexpected and unprecedented structure of the MoFe protein containing the Fe_7MoS_9(homocitrate) active site was determined crystallographically,[5] and then further in 2002, with a higher resolution crystal structure[6] that revealed the presence of a single atom, probably nonexchangeable nitrogen (N^c),[7,8] at the center of the active site cluster. The FeMo-cofactor (FeMo-co, $N^cFe_7MoS_9$(homocitrate)), shown in Figure 1, is linked to the protein via cysteine and histidine residues.[9]

Nitrogenases catalyze reaction 1. In broad terms, the sequence

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \quad (1)$$

of events for provision of electrons involves an ancillary Fe_4S_4 protein with two MgATP binding to the MoFe protein, hydrolysis of MgATP, transfer of one electron to the MoFe protein and, via a mediator P-cluster within it, to the FeMo-co active

(1) Peters, J. W.; Fisher, K.; Dean, D. R. *Annu. Rev. Microbiol.* **1995**, *49*, 335−366. Howard, J. B.; Rees, D. C. *Chem. Rev.* **1996**, *96*, 2965−2982. Smith, B. E. *Adv. Inorg. Chem.* **1999**, *47*, 159−218. Smith, B. E.; Durrant, M. C.; Fairhurst, S. A.; Gormal, C. A.; Gronberg, K. L. C.; Henderson, R. A.; Ibrahim, S. K.; Gall, T. L.; Pickett, C. J. *Coord. Chem. Rev.* **1999**, *185−186*, 669−687.
(2) Burgess, B. K.; Lowe, D. J. *Chem. Rev.* **1996**, *96*, 2983−3011.
(3) Christiansen, J.; Dean, D. R.; Seefeldt, L. C. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52*, 269−295.
(4) Igarashi, R. Y.; Seefeldt, L. C. *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, 351−384.
(5) Kim, J.; Rees, D. C. *Nature* **1992**, *360*, 553−560. Bolin, J. T.; Ronco, A. E.; Morgan, T. V.; Mortenson, L. E.; Xuong, N. H. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1078−1082.
(6) Einsle, O.; Tezcan, F. A.; Andrade, S. L. A.; Schmid, B.; Yoshida, M.; Howard, J. B.; Rees, D. C. *Science* **2002**, *297*, 1696−1700.
(7) Dance, I. *Chem. Commun.* **2003**, 324−325.
(8) Lee, H.-I.; Benton, P. M. C.; Laryukhin, M.; Igarashi, R. Y.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. *J. Am. Chem. Soc.* **2003**, *125*, 5604−5605. Lovell, T.; Liu, T.; Case, D. A.; Noodleman, L. *J. Am. Chem. Soc.* **2003**, *125*, 8377−8383. Hinnemann, B.; Norskov, J. K. *J. Am. Chem. Soc.* **2003**, *125*, 1466−1467. Hinnemann, B.; Norskov, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 3920−3927.
(9) The labeling of atoms, amino acids, and chains throughout this paper is that of the protein from *Azotobacter vinelandii*, Protein Database 1M1N.

10.1021/ja0481070 CCC: $27.50

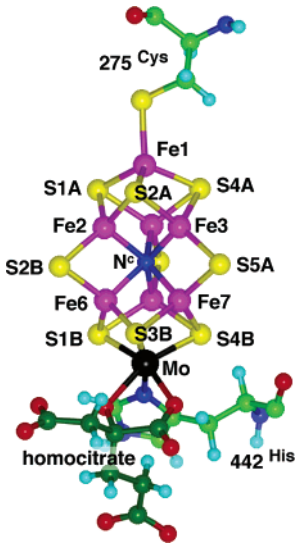


Figure 1. Structure of FeMo-cofactor, coordinated by 275^{Cys} at Fe1 and 442^{His} and homocitrate (C atoms darker green) at Mo. Atom labels are those of the crystal structure 1M1N: the unlabeled atoms at the back are (top down) Fe4, S3A, and Fe5. The color code is Mo black, Fe magenta, S yellow, C green, N blue, O red, and H cyan.

site.[4] Channels through the MoFe protein that could facilitate the ingress of gaseous substrates and egress of reactants have been described.[4,10]

Where and how might the substrates (which include alkynes, alkenes, and other small unsaturated molecules in addition to N_2) bind to the FeMo-co active site? In the resting state the Mo atom is six-coordinated, Fe1 is tetrahedrally coordinated, and the six central Fe atoms are four-coordinated, Fe(μ_3-S)$_2$-(μ-S)(N^c), in an inwardly trigonal pyramidal fashion toward N^c. The arguments for involvement of Mo (after partial dissociation of homocitrate) or for utilization of the central Fe atoms in the mechanism have been reviewed.[11] It is hard to avoid an overall view that the central Fe atoms, exposed for additional coordination, provide the reaction zone,[12] while the Fe1 and Mo ends are anchors to the surrounds, including the hydrophilic aqueous domain around homocitrate. Why would evolution construct the unusual central region of the active site if not to be part of a special reaction?

While research on the mechanism of nitrogenase is relatively well-informed about sequences of events,[2,13] there are some fundamental impediments to detection and spectroscopic characterization of intermediates. The trapping and further investigation of the protein with bound substrate or intermediates is complicated by the fact that protons are also obligatory reactants: reduction of the FeMo protein is a prerequisite for substrate binding, but reduced FeMo protein is oxidized back to the inactive resting state by ubiquitous protons. However, it has been possible to obtain spectroscopic data for the wild-type enzyme with bound inhibitor CO molecules,[14] and to obtain signals from freeze-quenched CS_2 molecules.[15]

In this context, recent experiments with modified proteins and with functionalized substrates have yielded new significant direct information. Dean and Seefeldt et al. have shown that the FeMo protein with α-70^{Val} substituted by the smaller alanine residue has activity with alkynes larger than ethyne. With this modified MoFe protein under turnover conditions, propargyl alcohol (HC≡C−CH_2OH) and propargyl amine are reduced and inhibit the reduction of N_2, C_2H_2, and protons. Most significantly, in this reaction a paramagnetic ($S = {}^1/_2$) intermediate can be trapped and characterized with EPR and ENDOR spectroscopy.[16−18] This led to a strong argument that the binding site is located on the Fe_4S_4 face comprised of iron atoms Fe2, Fe3, Fe6, and Fe7 (hereafter denoted the "Fe2,3,6,7 face") which is covered by the α-70 residue. The concept is that the α-70^{Ala} protein has a larger binding pocket on this face, enabling an increase in alkyne substrate size to propyne, 1-butyne, or functionalized (−OH, $-NH_2$) propyne. Detailed ENDOR spectroscopic analysis by Hoffman et al. showed that the observed species is allyl alcohol ($C(3)H_2$═$C(2)H$−$C(1)H_2OH$) bound to the face such that the two hydrogen atoms on C(3) are indistinguishable, leading to proposed η^2-coordination at one Fe atom.[18] Further investigation of the pH dependence of appearance of the $S = {}^1/_2$ intermediate for propargyl-OH and propargyl-NH_2, exploiting the fact that the pK_a of the imidazole of histidine is between those of propargyl-OH and propargyl-NH_2, resulted in the conclusion that the stabilization of the intermediate is aided by hydrogen bonding with α-195^{His}.[19]

The information accumulated in these experiments provides the best definition yet of the location of the binding of an alkyne and of its doubly hydrogenated alkene form. The objective of this paper is to interpret this information using theoretical and computational models, and to focus closely on the atomic detail of the location and geometry of intermediates bound to FeMo-co. The outcome is a quite specific description of the location and geometry of the interaction of propargyl substrates and allyl intermediates with FeMo-co, explaining the available data and supporting the earlier proposals, and leading to further detailed hypotheses about the mechanism of nitrogenase.

Setting the Coordination Scene

Before describing results, it is helpful to outline some of the key characteristics of the relevant coordination chemistry of FeMo-co with other substrates, as determined by density functional calculations. Coordination of a ligand L can occur in several ways at each central Fe (2−7), resulting in different coordination stereochemistry. The extremes for coordination at a single Fe are exo and endo. In the exo geometry (Figure 2a) the substrate is bound at Fe along extension of the N^c−Fe

(10) Durrant, M. C. *Biochem. J.* **2001**, *355*, 569−576.
(11) Seefeldt, L. C.; Dance, I. G.; Dean, D. R. *Biochemistry* **2004**, *43*, 1401−1409.
(12) Dance, I. G. *Aust. J. Chem.* **1994**, *47*, 979−990.
(13) Thorneley, R. N. F.; Lowe, D. J. In *Molybdenum enzymes*; Spiro, T. G., Ed.; Wiley-Interscience: New York, 1985; pp 221−284. Thorneley, R. N. F.; Lowe, D. J. *J. Biol. Inorg. Chem.* **1996**, *1*, 576−580.
(14) Pollock, R. C.; Lee, H. I.; Cameron, L. M.; DeRose, V. J.; Hales, B. J.; Orme-Johnson, W. H.; Hoffman, B. M. *J. Am. Chem. Soc.* **1995**, *117*, 8686−8687. Christie, P. D.; Lee, H. I.; Cameron, L. M.; Hales, B. J.; Orme-Johnson, W. H.; Hoffman, B. M. *J. Am. Chem. Soc.* **1996**, *118*, 8707−8709. Lee, H. I.; Hales, B. J.; Hoffman, B. M. *J. Am. Chem. Soc.* **1997**, *119*, 11395−11400. Lee, H. I.; Cameron, L. M.; Hales, B. J.; Hoffman, B. M. *J. Am. Chem. Soc.* **1997**, *119*, 10121−10126. Cameron, L. M.; Hales, B. J. *Biochemistry* **1998**, *37*, 9449−9456. Maskos, Z.; Hales, B. J. *J. Inorg. Biochem.* **2003**, *93*, 11−17.
(15) Ryle, M. J.; Lee, H.-I.; Seefeldt, L. C.; Hoffman, B. M. *Biochemistry* **2000**, *39*, 1114−1119.
(16) Christiansen, J.; Cash, V. L.; Seefeldt, L. C.; Dean, D. R. *J. Biol. Chem.* **2000**, *275*, 11459−11464.
(17) Mayer, S. M.; Niehaus, W. G.; Dean, D. R. *J. Chem. Soc., Dalton Trans.* **2002**, 802−807. Benton, P. M. C.; Laryukhin, M.; Mayer, S. M.; Hoffman, B. M.; Dean, D. R.; Seefeldt, L. C. *Biochemistry* **2003**, *42*, 9102−9109.
(18) Lee, H.-I.; Igarashi, R.; Laryukhin, M.; Doan, P. E.; Dos Santos, P. C.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. *J. Am. Chem. Soc.* **2004**, *126*, 9563−9569.
(19) Igarashi, R.; Dos Santos, P. C.; Niehaus, W. G.; Dance, I. G.; Dean, D. R.; Seefeldt, L. C. *J. Biol. Chem.* **2004**, *279*, 34770−34775.

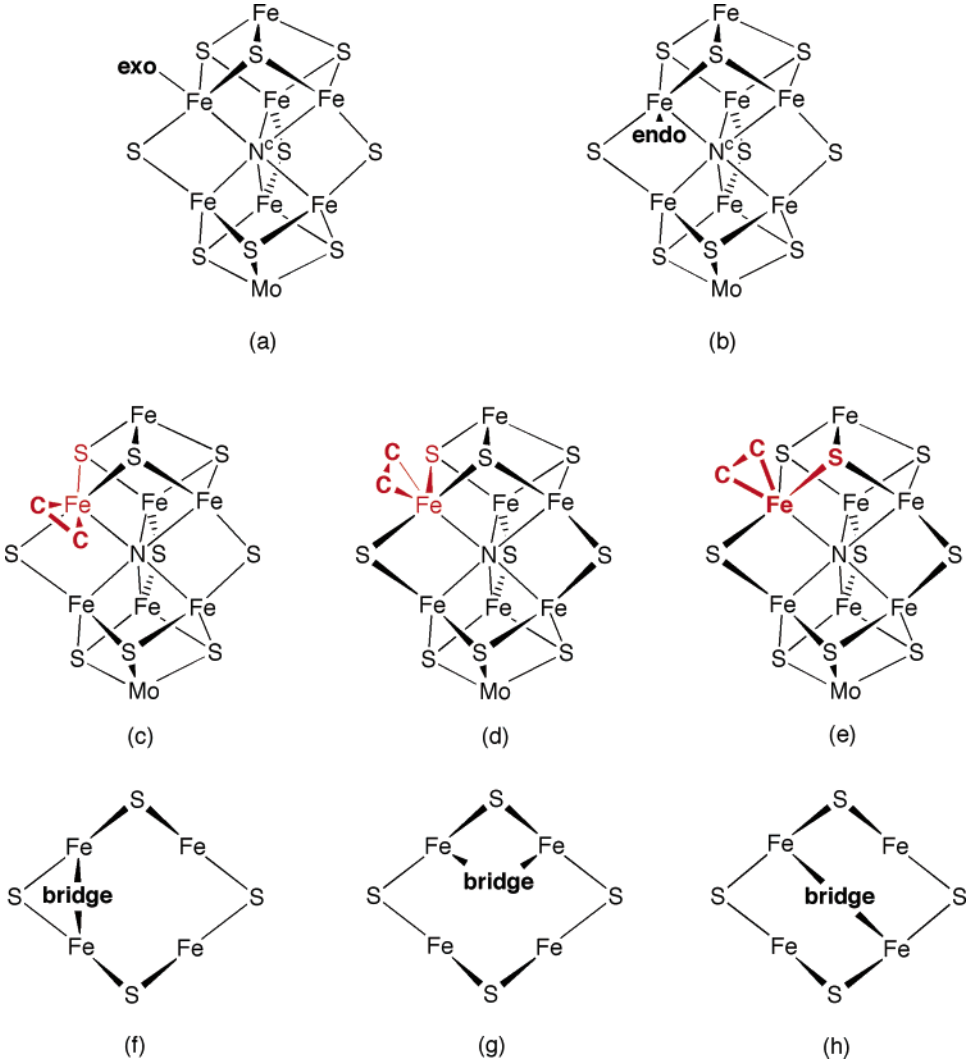
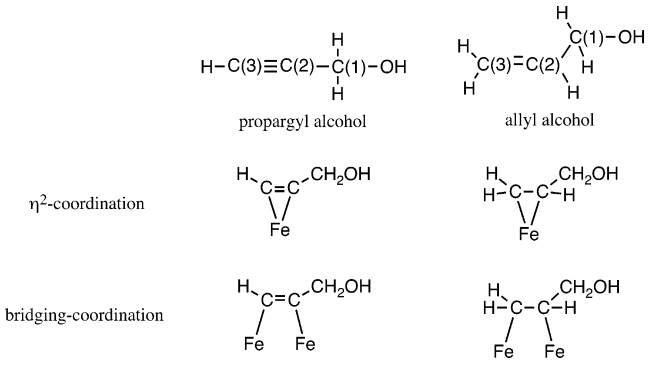


Figure 2. General coordination modes and isomers at one or more Fe atoms of the Fe_4S_4 face of FeMo-co. In conformers c−e, the red atoms are approximately coplanar.

bond: depending on the substrate and its bonding with Fe, the N^c−Fe distance can vary from ca. 2.2 Å to >3 Å, and thus the coordination stereochemistry of Fe can vary from trigonal bipyramidal to tetrahedral.[20] In the endo geometry (Figure 2b) the substrate is located approximately along an extension of a μ_3-S−Fe bond. There is a concomitant folding back of the μ-S atom involved, such that the set of four ligands $(\mu_3\text{-S})_2(\mu\text{-S})L$ are approximately square around Fe: the N^c−Fe bond, with variable length, can complete square-pyramidal stereochemistry at Fe. General characteristics of (single ligation) of Fe are that (1) the N^c−Fe distance is longer for exo coordination than for endo coordination, and (2) L can move on an arc between the endo and exo extremes. The central N^c atom appears to have a coordinative allosteric role, by which coordination at one Fe atom is transmitted to and influences coordinative abilities at another Fe atom.[20]

The possible coordination modes for the propargyl substrates and allyl intermediates are η^2 and bridging, as shown (for the alcohols) in Scheme 1. When the ligation is η^2 there is an additional variable, the conformational orientation of the CC bond relative to the $(\mu_3\text{-S})_2(\mu\text{-S})$ coordination. With endo

Scheme 1

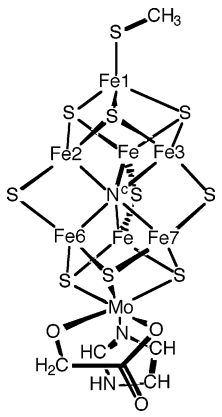


coordination only the transverse conformation of Figure 2c is possible, while with exo coordination the C−C bond can be aligned with either of the two μ_3-S atoms, as in Figure 2d,e. Bridging coordination can occur along edges of the Fe_4S_4 face in two ways (Figure 2f,g) or across a diagonal (Figure 2h).

The FeMo-co model **1** used in the density functional calculations includes the essential ligation of all atoms: $-SCH_3$ for 275^{Cys}; glycollate $-OCH_2COO-$ for homocitrate; and imidazole $C_3N_2H_4$ for 442^{His}. The resting state of the protein

(20) Dance, I. G., in preparation for publication, 2004.

corresponds to a charge of −3 for the $Fe_7MoS_9N^c(SCH_3)(OCH_2COO)(C_3N_2H_4)$ model, for reasons already described.[7]



1

Throughout this paper, the orientation shown for **1** defines directions (up/down, left/right, front/back) used in descriptions of geometry and location. The color code is Mo black, Fe magenta, S yellow, C green, N blue, O red, and H cyan; in some figures the C atoms of bound species are emphasized as black, and C atoms of the α-70 residue as brown. Hydrogen bonds are marked as cross-striped connectors.

Methods

Density functional calculations used the theory implemented in the program DMol of Delley,[21] which uses double numerical basis sets with polarization functions.[22] The calculations are all-electron (477 electrons for the unligated cluster) and spin-unrestricted, and use the blyp functional. This density functional methodology has been validated against a variety of metal sulfide clusters and related compounds[23] and has been applied extensively to the FeMo-co cluster and its properties.[20] Electron configurations were in general uncontrolled and (through use of Fermi smearing) allowed to adopt the orbital occupation that yields lowest energy: where the resulting spin state was other than $S = {}^1/_2$, the orbital occupancy was fixed to $S = {}^1/_2$, to be consistent with the experimental data. The electronic structure of **1** is characterized by multiple close-lying orbitals near the HOMO, and a HOMO−LUMO gap which was generally about 0.4 eV.

Since the objective is to map relevant parts of the energy surface for the binding of substrates, which can occur in many ways (see results), second derivative calculations were not undertaken, and the qualities of apparent local minima were assessed only by inspection of the behavior of the energy gradients during optimization.

Calculations of interactions with the surrounding protein included a large protein component comprised of 1032 amino acid residues, 1332 associated water molecules, and the P-cluster, together with FeMo-co. This was selected as all of chains A and B of 1M1N, together with residues 494−523 of chain D, and all water molecules within 4 Å of these atoms or FeMo-co. All hydrogen atoms were added, and energy-minimized by force-field methods (force-field cvff of Discover[24]), to optimize the protein and water hydrogen bonding. The P-cluster and the $N^cFe_7MoS_9O_2$ atoms of FeMo-co were fixed during all force-field calculations. The resulting optimized part-protein, denoted **Q**, had α-195^{His} protonated at both Nδ and Nε, to be consistent with the experimental data.[19] The procedure for testing the fit and interactions of a ligated FeMo-co model (**M**) in the α-70^{Ala} protein was to convert **Q** to α-70^{Ala}-**Q**, and then to superimpose **M** onto the FeMo-co cluster of α-70^{Ala}-**Q**. Because there is some geometrical change in the structure of **M** according to the nature of the ligation (see results), the least-squares superimposition was calculated using only the undistorted sections of **M** and FeMo-co, usually about 14 pairs of corresponding atoms: the rms deviation of these atom positions ranged from 0.15 to 0.25 Å. The original FeMo-co cluster was then deleted, leaving α-70^{Ala}-**Q** substituted by **M**. This construct of the α-70^{Ala} protein around ligated FeMo-co was examined for any conflicts between the ligand L and α-70^{Ala}, or between other parts of **M** and **Q**, and then force-field optimized keeping the $N^cFe_7MoS_9O_2$ + L atoms and the P-cluster fixed. A similar procedure was used to construct models with bound C_2H_2 or C_2H_4 inside the α-70^{Val} protein.

Results

My density functional (DF) calculations show that alkynes and alkenes can bind to FeMo-co in a variety of ways, involving one or two Fe atoms. The types of alkene binding that could meet the experimental criteria are illustrated generally in Figure 3, for allyl-OH. Structures **2** and **6** are η^2-Fe binding modes to one Fe atom, while the other five structures involve two Fe atoms with Fe−C−C−Fe connections: structures **23**, **26**, and **67** have the alkene bridging an edge of the Fe_4 quadrilateral, and **36** and **27** use the two diagonals.

The Binding of Allyl Alcohol. Details of relevant DF calculated structures for allyl-OH bound to **1** in the observed $S = {}^1/_2$ state are portrayed in Figure 4. The relatively invariant additional coordination of Fe1 and Mo is omitted in all pictures, which are viewed in the standard orientation of Figure 1.[25] Structure labels use the numbers of the Fe atoms involved. There is a strong tendency for the OH group to form a good O−H···S hydrogen bond to S2B, and this hydrogen bond influences the conformation of the CH_2OH end of the ligand, and positions the OH group to the region where it could accept a hydrogen bond from Nε−H of α-195^{His}. It is important to recognize that the FeMo-co core is plastic, and that ligation to any one or pair of iron atoms affects the geometry of the others. A general characteristic of FeMo-co is that elongation of one N^c−Fe bond has a relatively small energy penalty, as does variation of the stereochemistry at N^c: these distortions are evident in the pictorial representations of Figure 4.

Structures **2a**, **6a**, and **6b** have η^2-Fe(C═C) bonding in the exo position, which is characterized by pronounced elongation of Fe−N^c, and pseudo-tetrahedral coordination of Fe. Structures **6a** and **6b** show the rotational isomerism (Figure 2d,e) in which the C═C bond is aligned with one or the other of the two Fe−(μ_3-S) bonds. Structure **2b** represents a shift of the η^2-Fe(C═C) coordination around the exo ⇒ endo arc from **2a**, which decreases the N^c−Fe2 distance, changes the coordination stereochemistry, and strongly affects the position of the OH group. Similarly, structures **6a**, **6c**, **6d**, and **6e** are positions shifting around the exo ⇒ endo arc on Fe6, with consequences for the location of the CH_2OH atoms of allyl alcohol. Progression around this exo ⇒ endo arc also causes S2B to fold back as the S2B−Fe−S3B angle increases toward 180° concomitant

(21) Delley, B. *J. Chem. Phys.* **1990**, *92*, 508−517. Delley, B. In *Modern density functional theory: a tool for chemistry*; Seminario, J. M., Politzer, P., Eds.; Elsevier: Amsterdam, 1995; Vol. 2, pp 221−254. Delley, B. *J. Chem. Phys.* **2000**, *113*, 7756−7764.
(22) DMol3 version MS 3.0, 2003 (www.accelrys.com/mstudio/ms_modeling/dmol3.html).
(23) Dance, I. G.; Fisher, K. J. *Prog. Inorg. Chem.* **1994**, *41*, 637−803. Dance, I. G. In *Transition Metal Sulfur Chemistry: Biological and Industrial Significance*; Stiefel, E. I., Matsumoto, K., Eds.; American Chemical Society: Washington, DC, 1996; Vol. 653, pp 135−152. Dance, I. *J. Chem. Soc., Chem. Commun.* **1998**, 523−530.
(24) www.accelrys.com/insight/discover.html.
(25) Cartesian coordinates for all structures are available from the author.

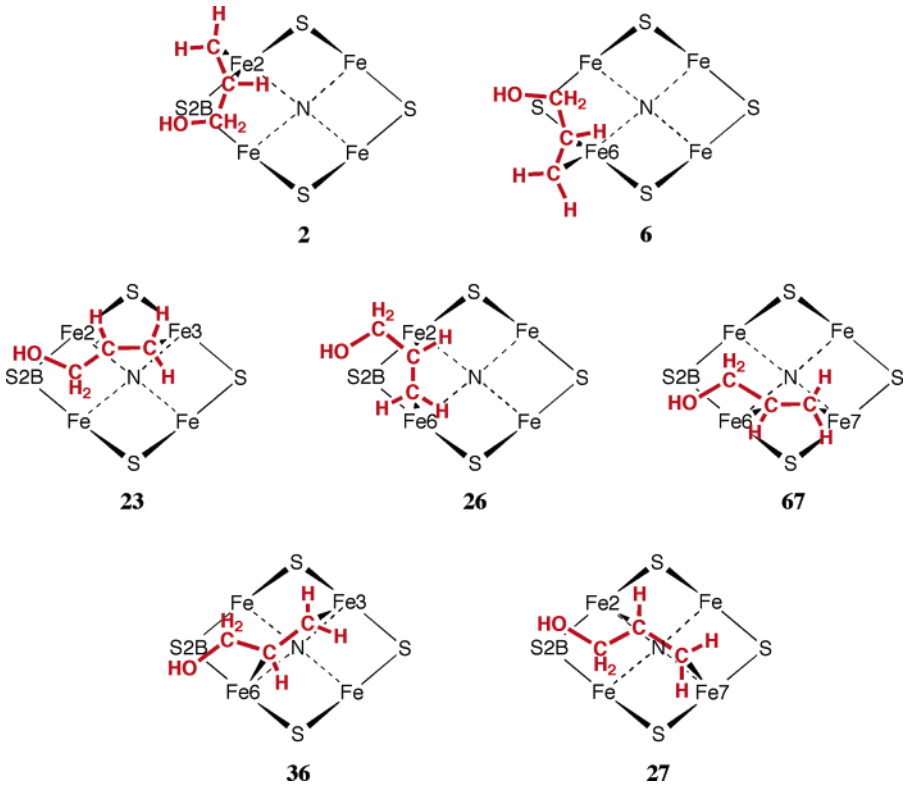


Figure 3. Relevant general structures for allyl alcohol bound to the Fe2,3,6,7 face of FeMo-cofactor. Structure labels use the numbers of the Fe atoms involved.

with the change in stereochemistry of Fe6 from pseudo-tetrahedral in **6a** toward square in **6e**.

Of the edge-bridging structures with Fe—C═C—Fe linkages, there are two conformers (**23a** and **23b**) which bridge the Fe2 and Fe3 atoms and are able to form an O—H···S2B hydrogen bond. Note that both Fe atoms have pseudo-square-pyramidal coordination, and that the μ_3-S atom S2A between Fe2 and Fe3 shifts upward. Structure **67** is a corresponding conformer for the Fe6—Fe7 edge. Atoms Fe2 and Fe6, connected instead by a μ-S atom, are bridged in structure **26**. The two possible structures in which a diagonal of the Fe_4 face is bridged are **36** and **27**: in both, the Fe_6 trigonal prism is substantially distorted in order that the two bridged Fe atoms be separated by ca. 3.3 Å. A characteristic of all bridged structures is a lengthening of the Fe—S bond trans to each C—Fe by ca. 0.2 Å: in the extreme case of structure **27**, energy minimization results in severance of the Fe2- -S1A bond, leaving pseudo-tetrahedral coordination of Fe2 by N^c, C(2) of allyl-OH, S2B, and S2A.

Note that structures **2b**, **6d**, **6e**, **23b**, **67**, **36**, and **27** have similar positions for the OH group of allyl-OH, and similar O—H···S2B hydrogen bonds.

Not all of the structures shown in Figure 4 are energy minima. The caption lists the relative energies of the structures portrayed. The η^2-Fe(C═C) structures on the exo—endo continuum have similar calculated energies on a relatively flat energy surface (evident also in the calculated gradients). The energy variation around the exo—endo arc is ca. 15 kcal mol^{-1}, in favor of the exo limit, while the edge-bridged structures are apparently local energy wells, calculated to be ca. 25—45 kcal mol^{-1} less stable than the η^2-Fe(C═C) structures.

With this cumulative knowledge of the coordination types for allyl-OH, and the relevant principles for geometry and energy, the key question to be considered is how they fit into the modified protein and whether they can satisfy the experimental criteria of hydrogen bonding with α-195^{His} and non-interference with other surrounding residues, especially α-70^{Ala}. This has been investigated by substitution of the DF-calculated structures for **1**—allyl-OH in place of FeMo-co in the modified α-70^{Ala} protein, examination of the hydrogen bonding and any conflicts, and then relaxation of the protein around the fixed DF structure to assess what changes might occur. The substitution was made by least-squares superimposition of about 14 corresponding Fe, Mo, or S atom pairs of FeMo-co and **1**—allyl-OH least affected by the additional coordination: the rms discrepancy in these atom positions in this substitution was ca. 0.2 Å, confirming the commensurability of the DF-calculated and experimental structures of the core. For the calculation of the fits of ligated FeMo-co into the FeMo protein, and subsequent relaxation of the protein, 1032 residues of 1M1N were included, together with 1332 water molecules included in this domain, with all hydrogen atoms explicitly included, and all hydrogen bonds optimized. This was designed to ensure that all secondary and tertiary aspects of the protein structure, and interactions of the protein with FeMo-co ligated by allyl-OH, would be included.

Some of the DF-calculated structures of Figure 4 clearly could not meet the experimental criteria, others could marginally, and a few fitted well. I will describe first the best fits, and then the problems with others, leading to the general principles that appear. The two principal geometric criteria are compatibility with the CH_3 group of α-70^{Ala}, and the hydrogen bond from N—H of α-195^{His} to O of allyl-OH; other general issues are the position of α-381^{Phe} affected by the folding back of S2B, and possible interference with α-96^{Arg} which can form hydrogen

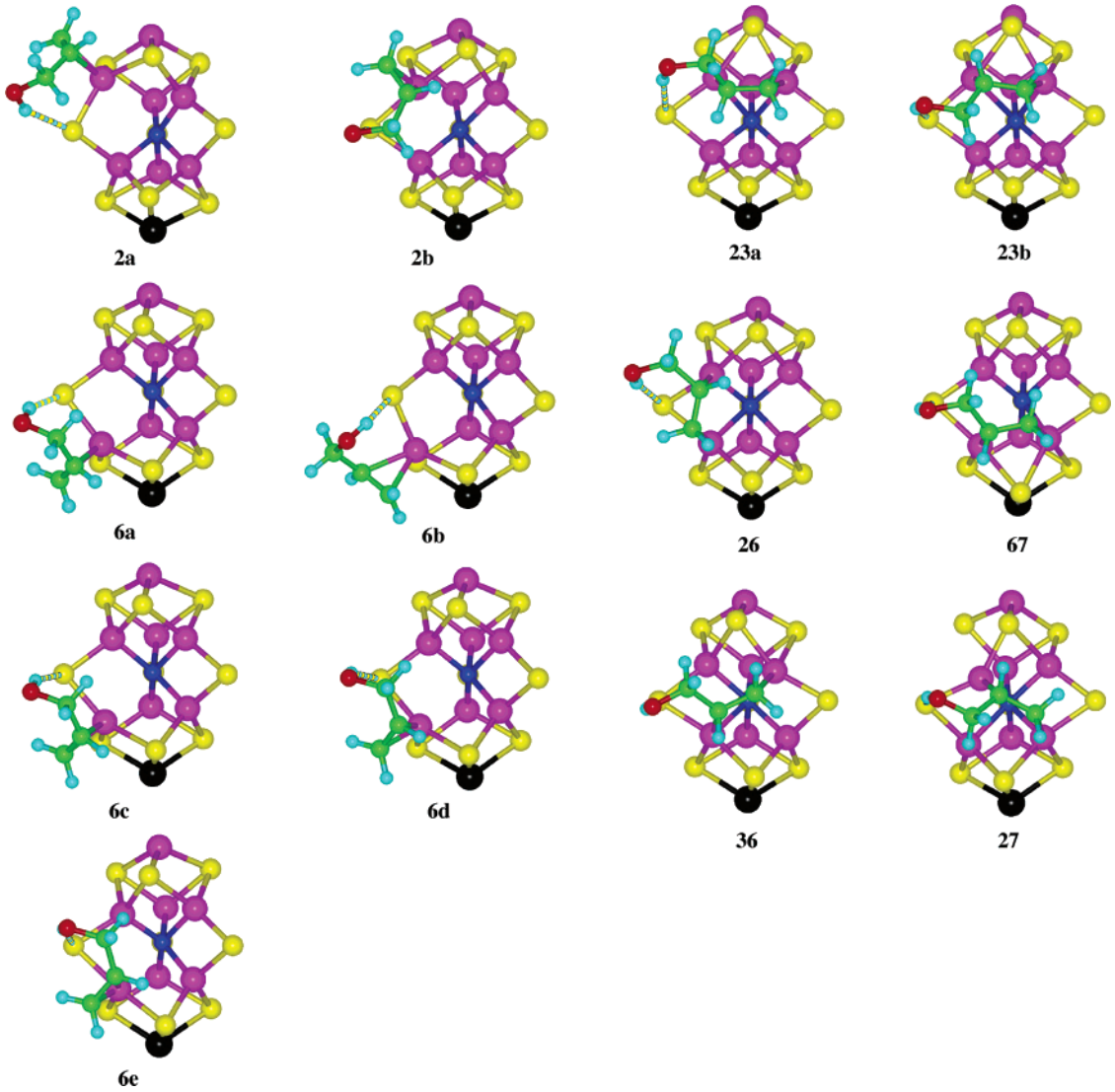


Figure 4. DF-calculated structures of allyl-OH bound to **1** (terminal ligands on Fe1 and Mo not drawn), all in the standard orientation of Figure 1. Structure labels use the numbers of the Fe atoms involved. Fe−N^c interactions longer than 2.5 Å are not drawn as bonds. The O−H···S2B hydrogen bonds are marked with yellow-cyan stripes. In **27**, the tight tetrahedral {N^cCS_2} coordination of Fe2 causes severance of the usual Fe2−S1A bond. Relative energies (kcal mol^{-1}) of the structures are **2a**, 10; **2b**, 11; **6a**, 0; **6b**, 6; **6c**, 2; **6d**, 8; **6e**, 22; **23a**, 37; **23b**, 38; **26**, 24; **67**, 45; **36**, 48; and **27**, 43.

bonds with S3B and S5A. In the following descriptions, the directions are referred to the standard FeMo-co orientation (Figure 1), looking directly at the Fe2,3,6,7 face, with Fe1−275^{Cys} at the top and Mo at the bottom.

The best fit of allyl-OH bound to FeMo-co in the α-70^{Ala} protein occurs for isomer **6d**, and is pictured in Figure 5, in comparison with the wild-type protein. The N−H···O and O−H···S2B hydrogen bonds are very well formed, and the methyl side chain of α-70^{Ala} is in almost the same position (relative to FeMo-co) as the Cβ part of α-70^{Val} in the wild-type protein. To accommodate the foldback of S2B, α-381^{Phe} also moved backward (not shown), without impediment.

When structures **6c** or **6e**, closely related to **6d**, are fitted into the protein and relaxed, the N−H···O hydrogen bond is poorer, being extended in **6c** and compressed in **6e**. This demonstrates an important principle, that α-195^{His} is a relatively constrained residue in this section of the protein, and is unable to move very far to maintain this hydrogen bond. There is no conflict between **6c** and the α-70^{Ala} side chain, but in **6e** the C(1)H_2 group of allyl-OH is shifted toward the methyl group of α-70^{Ala}, which moves down and right on relaxation, disfavoring **6e** as a possible structure for coordinated allyl-OH.

This interference between a C(1)−H atom of bound allyl-OH and the CH_3 side chain of α-70^{Ala} is a general factor, because, as illustrated in Figure 6, if OH is hydrogen bonded to S2B, and C(2) lies against the Fe_4 face because it and C(3) are bound to Fe atom(s) of the face, then the conformation of C(2)−C(1)H_2−OH must be such that one H atom protrudes approximately normal to the Fe_4 face. This protruding H atom is encircled in red in Figure 6. By comparison of the position of this protruding H atom with the position of the side chain of α-70, potential interference can be recognized: the other H atoms of bound allyl-OH protrude less.

How should such interference between protruding atoms of bound allyl-OH and the side chain of residue α-70 be assessed? Some insight comes from the way in which it is avoided during calculated relaxation of the protein. I first describe how any movement of α-70 is related to the local secondary structure. Amino acid α-70 is near the end of a short α-helix (from residues 62 to 72), which runs approximately laterally across

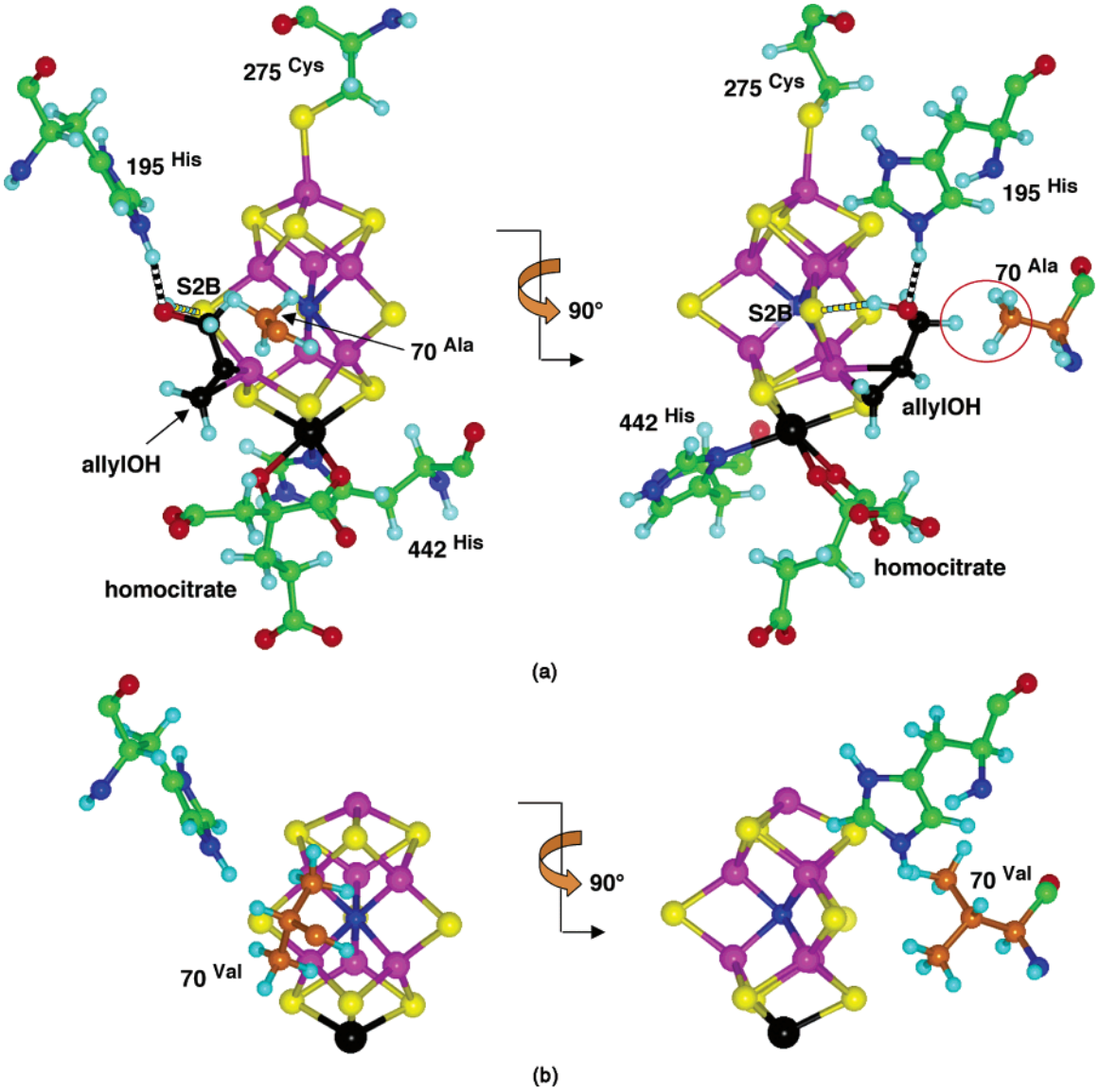


Figure 5. Partial representations of structure **6d** within the α-70^{Ala} protein (a) in comparison with the key sections of the wild-type α-70^{Val} protein (b). C atoms of allyl-OH are colored black, and the Cα, Cβ (and Cγ) atoms of 70$^{Ala\ (Val)}$ residue are colored brown; in the left views the NH and CO parts of residue 70 are omitted, for clarity. The N—H···O hydrogen bond is striped black and white, while the O—H···S2B hydrogen bond is striped yellow and cyan. Note the fold-back of S2B in the right view of (a), and that the side chain of residue 70 is similarly positioned over the face in the **6d**-α-70^{Ala} and wild-type α-70^{Val} proteins. The H···H contacts between allyl-OH and 70^{Ala} (enclosed in red) are 1.98, 2.06, and 2.86 Å.

the Fe2,3,6,7 face from left to right: therefore, left−right movement of α-70 involves displacement along the α-helical axis, up−down movement involves displacement around the helix, and movement away from the face involves displacement of the helix. During energy minimizations of the protein it appeared, qualitatively, that the energy gradient for movement of α-70^{Ala} away from the face was less than for lateral movements: away shifts of <0.5 Å appear to be feasible. The dynamics of protein structure variation during action, relative to the static crystal structure as the only quantitative structural data, are discussed at the end of this paper.

Considering the other DF structures for bound allyl-OH (Figure 3), exo-coordinated **2a** cannot form the expected hydrogen bonds, while the more endo-coordinated structure **2b** has an interference between a C(3)H atom and the face of the imidazole ring of α-195^{His} (see Figure 7). This is a fundamental impediment, in which one end of bound allyl-OH interferes with the hydrogen bond to the other end. Structures **6a** and **6b** cannot

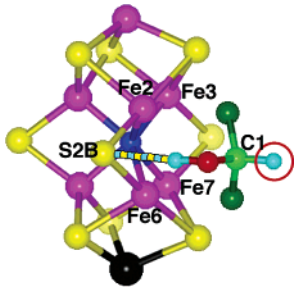


Figure 6. Generalized representation of the conformation of the HOC(1)H$_2$ section of allyl-OH bound to the Fe2,3,6,7 face of FeMo-co, such that an H—O···S2B hydrogen bond exists. The olive-green atoms represent the possible alternative locations of C(2) (bonded to Fe2 or Fe6) and H, while the relatively invariant H atom that protrudes normal to the Fe$_4$ face is encircled in red.

form acceptable hydrogen bonds with α-195^{His}; structure **23a** can be accommodated by α-70^{Ala}, but does not form an acceptable hydrogen bond with α-195^{His}; structure **23b** forms

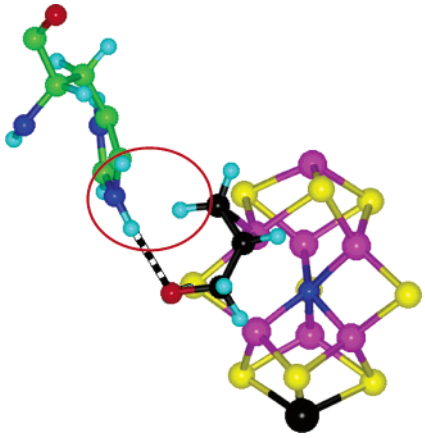
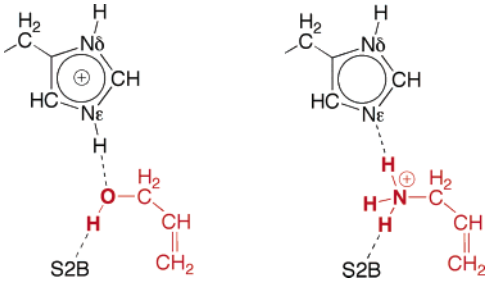


Figure 7. Part of structure **2b** in the protein, emphasizing the distorting steric conflict of allyl-OH with α-195^{His} and its hydrogen bond.

Scheme 2



a good hydrogen bond, but is obstructed by α-70^{Ala}; structure **26** has no interference with α-70^{Ala}, but the high position of OH does not permit an acceptable hydrogen bond with α-195^{His}; structure **67** forms a good hydrogen bond, but the protruding H atom of allyl-OH conflicts with α-70^{Ala}, and in addition the C(3)H_2 end of allyl-OH causes displacement (not shown) of α-96^{Arg}, an important residue; **36** has a good hydrogen bond, some interference with α-70^{Ala}, and little affect on α-96^{Arg}; structure **27** forms a good hydrogen bond with α-195^{His}, but interferes with α-70^{Ala} and α-96^{Arg}.

The conclusions about the coordination of allyl-OH to FeMo-co in the α-70^{Ala} protein are (1) the geometry of **6d**, intermediate between exo and endo on Fe6, is best; (2) the corresponding positions on Fe2 cause conflict with the imidazole face of α-195^{His} and are unlikely; (3) the bridging coordination is feasible for both **23b** and **36**, although bridging Fe—C═C—Fe coordination is less stable than η^2-Fe(C═C) coordination; and (4) the other structures are defective and unlikely.

The Binding of Allyl Amine and Its Conjugate Acid. The DF-calculated structures for allyl-NH_2 and allyl-NH_3^+ bonding to one or two Fe atoms of the Fe2,3,6,7 face are in general similar to those already described (Figure 4) for allyl-OH. For allyl-NH_2 there is one N—H···S2B hydrogen bond, and the other N—H can have either of two configurations. When the amine is protonated and conformed to a N—H···S2B hydrogen bond, the density functional energy minimization of **1** + allyl-NH_3^+ led to transfer of H along the hydrogen bond from N to S2B, i.e., N—H···S2B became N···H—S2B. This apparent greater basicity of S2B than of coordinated allyl-NH_2 may be an artifact of the absence of surrounding protein in the DF calculation. For the testing of coordinated structures in the protein these possibilities are sterically equivalent, and the N—H···S2B isomer was used.

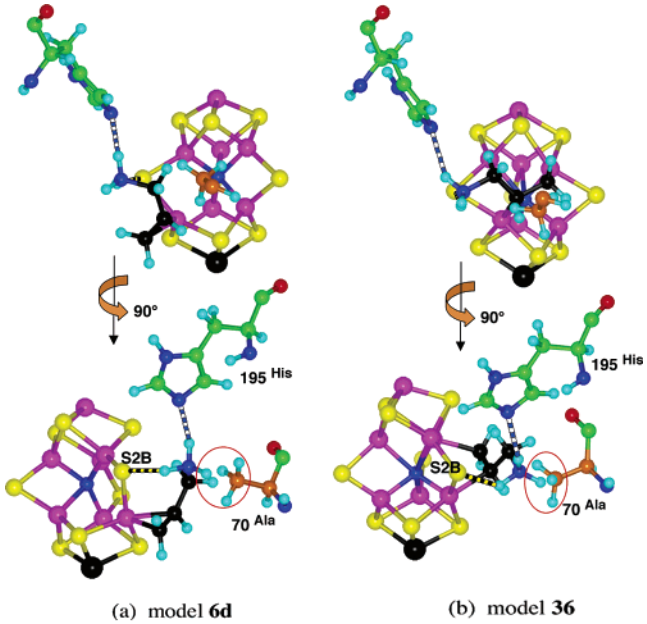


Figure 8. (a) Two views of structure **6d** for allyl-NH_3^+ coordinated to FeMo-co in the α-70^{Ala} protein. The C atoms of allyl-NH_3^+ are black, Cα and Cβ of 70^{Ala} are brown, the N—H···S2B hydrogen bond (2.1 Å, 156°) is black/yellow, and the N—H···Nϵ hydrogen bond (2.0 Å, 164°) is blue/white. The closest H···H contacts between allyl-NH_3^+ and 70^{Ala} (red enclosure) are 1.98 and 2.03 Å. (b) The corresponding representations for structure **36**. The closest H···H contact between allyl-NH_3^+ and 70^{Ala} (red enclosure) is 2.13 Å.

The experimental data indicate that in the observable stabilized species, the hydrogen bonding between allyl-NH_3^+ and α-195^{His} is the reverse of that for allyl-OH, as shown in Scheme 2. The DF-calculated structures for coordinated allyl-NH_3^+ were therefore evaluated in the protein for their ability to meet this criterion. Most of the possibilities were discounted, due to their interference with α-70^{Ala} or very poor hydrogen bonding with α-195^{His}, or both. The best possibility is coordination structure **6d** (see Figure 8a), which has two good hydrogen bonds and has α-70^{Ala} within ca. 0.3 Å of its location in the wild-type protein. As with coordination of allyl-OH, movement of η^2-bound allyl-NH_3^+ around the exo—endo arc of coordination positions, to geometry **6e** or **6c** or further, introduces conflicts with the protein. The next best structure in the protein is **36** (Figure 8b), which has poorer hydrogen bonds and a larger displacement (ca 0.9 Å for Cβ) of α-70^{Ala}. The other seven DF structures for allyl-NH_2 or allyl-NH_3^+ bound to FeMo-co fitted badly into the protein, and are unlikely possibilities.

Thus, the modeling of both the allyl alcohol and allyl amine systems in the α-70^{Ala} protein leads to the conclusion that the most probable coordination geometry for the observed species is **6d**, which is consistent with all experimental data. In the ENDOR spectra of bound allyl-OH, the two hydrogen atoms on C(3) are indistinguishable,[18] and this also matches the calculated model **6d**, in which the immediate coordination environments of these two hydrogen atoms are almost identical (see Figure 9). The DF-calculated relative energies also show that the η^2-coordination of allyl alcohol or allyl amine to one Fe atom is more favorable than the bridging of two Fe atoms.

The NH_3^+ group of bound allyl-NH_3^+ in the proposed model occupies space (on the left side of the Fe2,3,6,7 face in the standard front view) in contact with α-66^{Gly}, which is adjacent to α-70 on the same side of the α-helix containing them both.

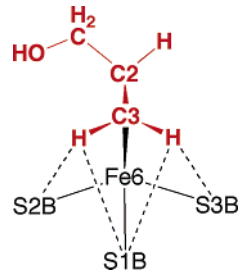


Figure 9. Close similarity of environments of the two H's on C(3) of allyl-OH in structure **6d**: the distances H−S2B, H−S3B are 2.9, 3.1 Å; the distances H- -S1B are 3.7, 3.8 Å.

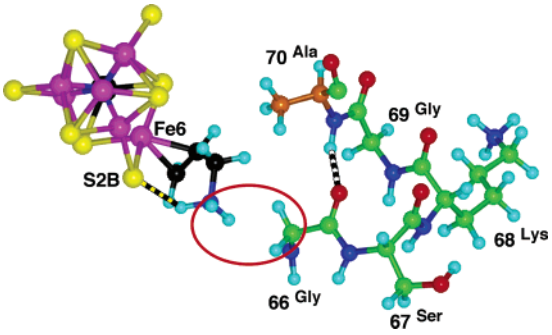


Figure 10. Part of model **6d** for allyl-NH_3^+ in the α-70^{Ala} protein, viewed from the top along the Fe1- -Mo axis of FeMo-co: atom coloration as in previous figures. The contact between the $-NH_3^+$ group and α-66^{Gly} is enclosed in red.

This is illustrated in Figure 10 (viewed along the Fe1- -Mo axis of FeMo-co), which shows how the chain circles from α-70^{Ala} on one side of bound allyl-NH_3^+ to α-66^{Gly} on the other, $-NH_3^+$, side. Overall this is a favorable juxtaposition of coordinated intermediate and local protein, and it also suggests that modification of residue α-66, increasing the volume of the side chain, could interfere with coordination of allyl-NH_3^+. This is a testable prediction.

There are still two questions. (1) Why are the propargyl alcohol or amine substrates used in these experiments not also observed bound to FeMo-co? (2) Are the preferred location and geometry for binding of the functionalized alkyne substrate and alkene intermediate in the α-70^{Ala} protein consistent with a similar location and geometry for interaction with ethyne and ethene in the wild-type α-70^{Val} protein? Modeling in response to these questions is described next.

Interactions of Propargyl Alcohol/Amine with FeMo-co. The EPR and ENDOR measurements have detected bound allyl alcohol or amine, without evidence for the coordinated propargyl reactants. The trapped species are late in the reaction, and are in fact the major alkene products.[17] It might be expected that the alkyne precursors would coordinate to Fe in a similar manner, and be similarly trapped.

I have explored (by DF methods) the coordination locations and modes for propargyl-OH and propargyl-NH_2/propargyl-NH_3^+, and Figure 11 illustrates the variety of structures for these coordinated alkynes.[26] Not all are energy minima, but these structures are energetically accessible: a full range of structures for bound alkynes is included in Figure 11 because they could be relevant to other experimental data not considered here. The sequences **P2a**−**P2b**−**P2c** and **P6a**−**P6b**−**P6c**−**P6d** are progressions around the exo ⟹ endo coordination arc, with concomitant variation in Fe coordination stereochemistry and FeMo-co geometry, as for bound alkenes. Note (Figure 11) that there are some reversals of position of the functionalized end of the alkynes, because the isomers portrayed are those that might match the experimental criteria of hydrogen bonding with α-195^{His} and nonconflict with α-70^{Ala}.

The propargyl structures were substituted (*in silico*) into the α-70^{Ala} protein, followed by calculated relaxation of the protein as before, and examination of the relationships with α-70^{Ala} and α-195^{His} in particular. There is a basic stereochemical difference between the propargyl molecules which are extended and the allyl molecules which are bent, causing the location of the OH/NH_3^+ group relative to the Fe(C_2)(Fe) coordination section to be different, and therefore the hydrogen-bonding capabilities of the bound propargyl-OH/NH_3^+ substrates are different from those of the allyl analogues. Because propargyl-NH_3^+ has less conformational freedom than allyl-NH_3^+, it was found that organization of a good (propargyl)N−H···N(195^{His}) hydrogen bond was generally not possible. The closest structure found for bound propargyl-NH_3^+ in α-70^{Ala} protein is **P6b**, depicted in Figure 12. Note that the imidazole ring of 195^{His} is forced upward (relative to protein structure 1M1N), degrading the quality of the N−H···Nϵ hydrogen bond.

This most likely position and coordination mode for the interaction of the propargyl reactants with FeMo-co, that is η^2 coordination on Fe6 in a position between exo and endo, is the same as that determined for the corresponding allyl intermediates. This reinforces the conclusions for both, because it is expected that the addition of two hydrogen atoms to propargyl reactant to yield allyl product would not involve large changes in coordination. I propose that the reaction using functionalized propargyl substrates in the α-70^{Ala} enzyme proceeds with weak binding and activation of the substrates as in Figure 12, followed by hydrogenation to functionalized allyl species which, while still bound as in Figure 8a, are conformationally able to form stronger hydrogen bonds with 195^{His}. This differentiation of hydrogen bonding suggests reaction sequences in which the steady-state concentrations of the bound allyl species are larger than those of the bound propargyl species, which is consistent with the spectroscopic observation only of the allyl species.

Could Ethyne React at the Same Location? The hypothesis of the experiments that led to this computatonal investigation—that residue α-70 encloses a site active for the binding and hydrogenation of substrates—is supported by the result that modification of residue α-70 from valine to the smaller alanine increases the volume available to substrates, and increases the size of reducible substrates from ethyne to the larger alkynes propyne and propargyl alcohol or amine. Implicit in this is that ethyne can bind and be hydrogenated at the same site in the wild-type α-70^{Val} protein (it is believed that there is more than one binding site/geometry for ethyne[11,27]). Therefore, I have investigated the binding of C_2H_2 and C_2H_4 in the newly proposed manner on Fe6 in the wild-type protein. Figure 13 shows the results for two coordination modes for C_2H_2 along the exo−endo arc, and one bonding position for C_2H_4, in the

(26) Some conformations of propargyl-NH_3^+ on **1** are such that energy minimization leads to dissociation of NH_3 and binding of delocalized $CH_2{=}C{=}CH^+$.

(27) Benton, P. M. C.; Christiansen, J.; Dean, D. R.; Seefeldt, L. C. *J. Am. Chem. Soc.* **2001**, *123*, 1822−1827. Han, J.; Newton, W. E. *Biochemistry* **2004**, *43*, 2947−2956.

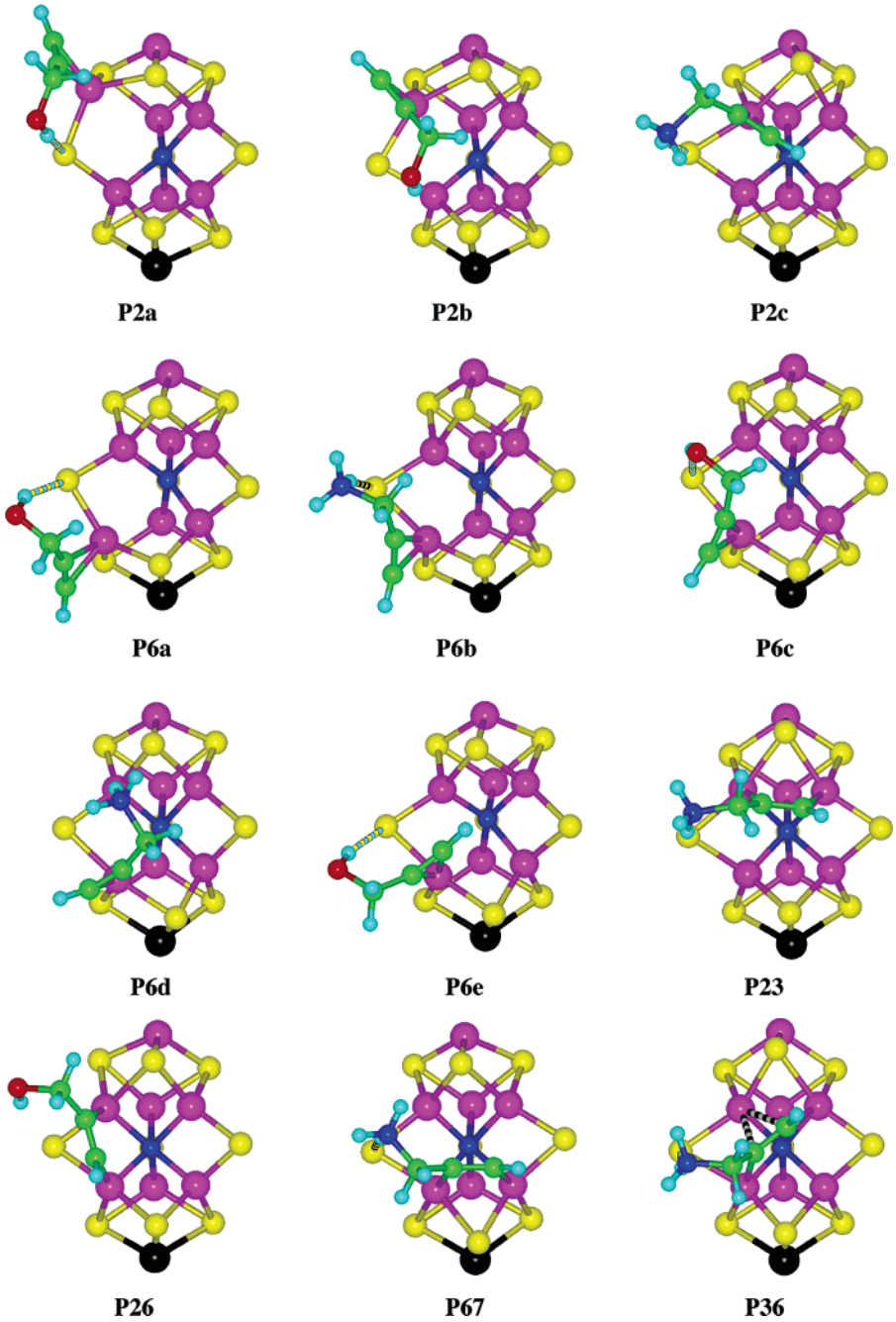


Figure 11. Calculated (DF) structures for propargyl-OH or propargyl-NH_3^+ bound to **1** (terminal ligands at Fe1 and Mo not shown). Each structure type is illustrated by only one or the other of propargyl-OH or propargyl-NH_3^+. In **P36**, the primary Fe−C bonds are to Fe6 and Fe3, but there are two longer (by ca. 0.3 Å) bonds to Fe2, marked with black and white stripes.

α-70^{Val} protein. Both molecules can bind η^2 to Fe6. It appears that acetylene can bind with more flexibility than can its hydrogenated product, ethylene, which is tightly constrained.

Discussion

In summary, this research has evaluated (by reliable theoretical methods) the possible structures for the binding of propargyl and allyl alcohols/amines to the FeMo-cofactor of the α-70^{Val} FeMo protein which are consistent with the recent experimental results from Seefeldt, Dean, Hoffman, et al.[16−19] A specific result ensues. The η^2-coordination of alkyne and alkene most probably occurs at Fe6 of the FeMo-cofactor, in a position (**6d**) that is intermediate between the exo and endo coordination extremes at Fe6. This model accounts for (1) the steric influence of the α-70 residue, (2) the crucial hydrogen bonding with Nϵ of α-195^{His}, (3) the spectroscopic symmetry of the allyl alcohol intermediate, and (4) the preferential stabilization of allyl alcohol/amine relative to propargyl alcohol/amine. The bridging atom S2B has an important hydrogen-bond-accepting role for these substrates. While structure **6d** is considered as most likely, alternative bridging structures **23b** and **36** may be possible if there can be larger movement of α-70 and its neighboring residues in the functioning protein. However, since the steric difference between the two functioning alternatives, (a) α-70^{Val} protein plus ethyne substrate and (b) α-70^{Ala} protein plus propyne or 1-butyne or propargyl substrates, is just one or two

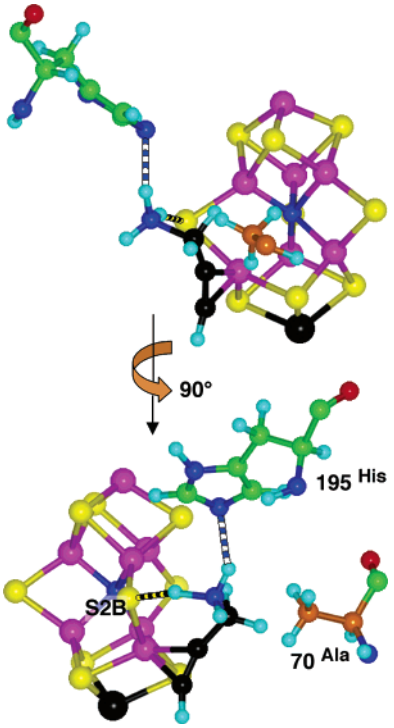


Figure 12. Part of the best model (**P6b**) for propargyl-NH_3^+ in the α-70^{Ala} protein. Note the upward distortive folding of 195^{His}, relative to the structures in Figure 8.

methyl groups, it is likely that there is little variability of protein structure in this region.

The DF calculations have shown a variety of ways in which alkynes and alkenes can bind to the relevant Fe2,3,6,7 face. Some of these, other than the modes favored for the propargyl and allyl molecules, could be relevant for other aspects of the reactivity of ethyne, such as alternative sites proposed to account for cis and trans hydrogenation of ethyne. These structures are also expected to inform more detailed postulates for the structures of the bound intermediates in later stages of the reduction of N_2, especially diazene HNNH and hydrazine H_2NNH_2. The complex enzyme−kinetic relationships between ethyne and dinitrogen[3,4] attest to the relevance of these alkyne binding sites to the main issue of the mechanism of dinitrogen reduction. Further modeling of these aspects, together with the inhibition by CO (for which there is spectroscopic characterization), is in progress.

The fitting of ligated FeMo-co into the protein undertaken here has involved interactions with α-381^{Phe}, α-96^{Arg}, α-66^{Gly}, and α-191^{Gln}, in addition to the key residues α-70 (with steric influence) and α-195^{His} (providing the hydrogen bonding). Many experimental investigations have uncovered influences by the residues α-69, α-70, α-96, α-191, and α-195.[3,4] Residue α-381^{Phe} is located behind S2B, which folds back as part of the changes in coordination stereochemistry at Fe2 or Fe6 or both. However, there does not appear to be any impediment to this shift of the α-381^{Phe} side chain. Residue α-66 is next to α-70 in the α-helix that contains them both, and therefore is close to coordinated substrates and intermediates. In the wild-type protein, α-66 is glycine, and it is predicted that an increase in the side chain volume at α-66 could diminish enzyme activity with small-molecule substrates.

This leads to issues of protein dynamics. The crystallographic structural information is from static FeMo protein in inactive

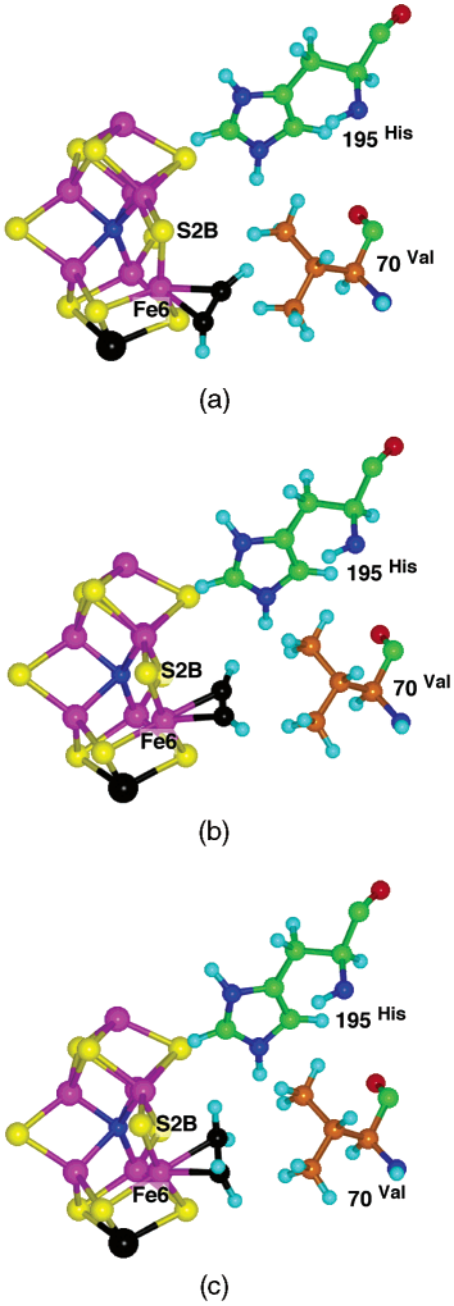


Figure 13. Models of DF-calculated structures of η^2-C_2H_2 (a, b) and η^2-C_2H_4 (c) bound to Fe6 in the wild-type α-70^{Val} protein.

form. A key question for investigation of mechanism is the variability and the dynamics of protein structure. While the current investigation has not addressed this specifically, it has explicitly included all hydrogen atoms and optimized hydrogen bonding and protein structure with a reliable force field: a number of observations are worth recording. The influential residues are contained in two α-helices that touch each other. One helix, containing α-66^{Gly} and α-70^{Val}, lies across the key Fe2,3,6,7 face, while the other, containing α-191^{Gln} and α-195^{His}, lies to the side, near S2B. This region and the proposed substrate/intermediate binding region are notably water-free, and there is limited hydrogen bonding, so the two key α-helices appear to be able to move independently. This deserves further investigation. The remarkably aqueous domain surrounding homocitrate is nearby, and these water molecules and their

hydrogen bonds seem to be mobile. The key residue α-195^{His} was relatively immobile in my models, and is associated via Nδ with only one water molecule (number 164 in 1M1N), which is absent in the crystal structure of the α-195^{Gln} mutant of the FeMo protein.[28] Residue α-195^{His} appears to be a hydrogen-bonding anchor, consistent with its suggested role of stabilizing hydrogenated intermediates in the reduction of dinitrogen. Finally, α-96^{Arg} showed some movement in my calculations, with variable hydrogen bonding to S atoms of FeMo-co, suggesting that it could be part of the proton delivery system.

(28) Sorlie, M.; Christiansen, J.; Lemon, B. J.; Peters, J. W.; Dean, D. R. *Biochemistry* **2001**, *40*, 1540−1549.

Igarashi and Seefeldt[4] identified a gas substrate channel through the protein from its surface to FeMo-co. When a CPK graphic of the protein is viewed along this channel, the bound intermediate (in positions **6d** or **23b** or **36**) is just visible, further supporting this as the site of nitrogenase activity.

Acknowledgment. This research is supported by the Australian Research Council, the University of NSW, and the Australian Centre for Advanced Computing and Communication. I thank Lance Seefeldt and Dennis Dean for provision of results prior to publication, and valuable discussions.

JA0481070