Determination of Noncovalent Docking by Infrared Spectroscopy of Cold Gas-Phase Complexes

Etienne Garand, *et al.*

Science **335**, 694 (2012);
DOI: 10.1126/science.1214948

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

29. A. P. R. Johnston, G. K. Such, F. Caruso, *Angew. Chem. Int. Ed.* **49**, 2664 (2010).
30. W. S. Jeon, H.-J. Kim, C. Lee, K. Kim, *Chem. Commun. (Camb.)* **38**, 1828 (2002).
31. S. Abalde-Cela *et al.*, *Soft Matter* **7**, 1321 (2011).
32. H. Hwang, S.-H. Kim, S.-M. Yang, *Lab Chip* **11**, 87 (2011).
33. W.-F. Dong, G. B. Sukhorukov, H. Möhwald, *Phys. Chem. Chem. Phys.* **5**, 3003 (2003).
34. S. Mahajan *et al.*, *Phys. Chem. Chem. Phys.* **12**, 10429 (2010).
35. J. R. Anema, J.-F. Li, Z.-L. Yang, B. Ren, Z.-Q. Tian, *Annu. Rev. Anal. Chem.* **4**, 129 (2011).

Acknowledgments: This work was supported by Engineering and Physical Sciences Research Council Basic Technology Translational grant EP/H046593/1, a European Research Council Starting Investigator Grant (ASPiRe) ERC-2009-StG-240629, and a European Union NanoSci-E+ (CUBiHOLE) grant EP/H007024/1. SERS experiments were performed with the assistance of S. Mahajan. The DH5α *Escherichia coli* strain was a kind gift from M. Welch and S. Bowden from the Department of Biochemistry, University of Cambridge. The naphthol-containing polymer was a kind gift from E. Appel from the Department of Chemistry, University of Cambridge. The authors thank X. Liu, C. A. Smith, S. Mahajan, J. J. Baumberg, and W. T. S. Huck for the helpful discussions. J.Z., R.J.C., O.A.S., and C.A. have filed a provisional application for a GB patent on the fabrication process.

Supporting Online Material

www.sciencemag.org/cgi/content/full/335/6069/690/DC1
Materials and Methods
Figs. S1 to S3
References (*36*–*40*)

18 October 2011; accepted 18 January 2012
10.1126/science.1215416

Determination of Noncovalent Docking by Infrared Spectroscopy of Cold Gas-Phase Complexes

Etienne Garand,[1] Michael Z. Kamrath,[1] Peter A. Jordan,[1] Arron B. Wolk,[1] Christopher M. Leavitt,[1] Anne B. McCoy,[2] Scott J. Miller,[1] Mark A. Johnson[1]*

Multidentate, noncovalent interactions between small molecules and biopolymer fragments are central to processes ranging from drug action to selective catalysis. We present a versatile and sensitive spectroscopic probe of functional groups engaged in hydrogen bonding in such contexts. This involves measurement of the frequency changes in specific covalent bonds upon complex formation, information drawn from otherwise transient complexes that have been extracted from solution and conformationally frozen near 10 kelvin in gas-phase clusters. Resonances closely associated with individual oscillators are easily identified through site-specific isotopic labeling, as demonstrated by application of the method to an archetypal system involving a synthetic tripeptide known to bind biaryl substrates through tailored hydrogen bonding to catalyze their asymmetric bromination. With such data, calculations readily converge on the plausible operative structures in otherwise computationally prohibitive, high-dimensionality landscapes.

A central strategy in both contemporary drug discovery and chemical synthesis exploits multiple weak or reversible hydrogen (H)–bonding interactions among embedded functional groups to achieve selective docking arrangements between complex yet flexible molecular architectures. Diverse examples include the peptidic interaction motifs underlying antibiotic activity (*1*) and asymmetric catalysis (*2*–*5*), as well as, in a larger sense, the field of supramolecular chemistry (*6*). Although simulation plays an increasingly important role in guiding these efforts (*7*–*10*), there is a clear demand for experimental methods that can probe the docking geometries of the noncovalently bonded partners. Many of the key attachment points involve complementary H-bonding between C=O, N-H, and O-H groups, with increasing evidence emerging for the special role that charged motifs (e.g., involving $-NH_3^+$ and $-CO_2^-$ interactions with neutral partners) play in enhancing noncovalent binding (*3*, *4*, *11*). In favorable cases, x-ray crystal structures, nuclear magnetic resonance, or two-dimensional vibrational spectroscopies (*12*–*14*) can yield this information, but more often the key species occur as short-lived transients that are not amenable to these approaches. Here, we demonstrate how a cryogenic vibrational spectroscopic method, carried out after the complexes are preformed in solution and then isolated in the gas phase, can be used to identify the specific C=O and N-H groups that link a small molecule to a polypeptide in a mass-selective variation of site-specific, isotope edited spectroscopy (*12*, *15*–*17*).

We focused on a protonated tripeptide catalyst, and a biaryl substrate molecule (Fig. 1), where the noncovalent docking complex is a positively charged variation of the intermediate invoked to rationalize stereoselective bromination of the substrate (*2*, *18*). The complex thus provides an archetypal example wherein multidentate interactions were designed to provide highly selective docking of a small molecule in the context of a flexible catalytic scaffold. The operative attachment motifs involve H-bonding; O-H groups on the substrate (**S**) and N-H groups on the catalyst (**C**) constitute the donors, and the O-H, =O, and –N: sites on both species can act as potential acceptors. Four important functional groups are denoted by Greek letters (α,β,γ,δ) in Fig. 1 for reference. Our goal was to elucidate experimentally which specific acceptor-donor pairs might be active in the complex formation. Structural characterization of the complex involved first establishing the spectral signatures of the intramolecular H-bonds that fold the isolated tripeptide catalyst. The sensitivity of these modes to H-bonding interactions was revealed by following how the H-bonding pattern changed when the excess proton, which dominates the H-bonded interactions in **C**, was replaced by the more mobile Na^+ ion that binds primarily through electrostatic interactions. We then determined which bonds are directly involved in a binding site that supports the docking arrangement with the biaryl substrate.

The key advance enabling vibrational spectroscopic characterization of these complexes is the recently developed ability to extract them from solution by electrospray ionization and quench them into stable local configurations by cryogenic ion processing (*19*–*21*). In this approach, the ions are first tagged with weakly bound hydrogen molecules (fig. S1) such that single-photon (linear) vibrational action spectra can be obtained by following the photoinduced mass loss accompanying excitation of vibrational resonances. In the present case, D_2 tags were used rather than H_2 because their larger molecular masses helped separate the parent and fragment signals. Such infrared (IR) spectra are obtained for mass-selected species with an instrument that integrates mass spectrometry with photofragmentation (*19*, *22*, *23*). This method is generally applicable, and yields spectra of complex ions and ion-solvent clusters cooled to around 10 K (*22*–*24*), which display intrinsic linewidths as low as 6 cm^{-1} in both the fingerprint region as well as the higher-frequency X-H stretches. Because the power requirements of the excitation source are very modest (compared to the more commonly used infrared multiple photon dissociation), the technique can be implemented with table-top laser sources (*19*, *20*), and the resulting linear absorption spectra are readily compared with calculated patterns of vibrational fundamentals.

The challenges confronting structural analysis of noncovalent interactions with IR spectroscopy are immediately evident in Fig. 1, which compares the D_2 predissociation spectrum of **C** with that of the **C-S** complex. Although both species exhibit many sharp transitions scattered throughout the mid-IR spectral region, the overall patterns are quite similar in the signature amide I region (mostly C=O stretches from 1600 to 1800 cm^{-1}) and the amide A region (largely N-H stretches from 3300 to 3500 cm^{-1}). Nonetheless, a few coarse features of the intermolecular binding ar-

[1]Sterling Chemistry Laboratory, Yale University, Post Office Box 208107, New Haven, CT 06520, USA. [2]Department of Chemistry, The Ohio State University, Columbus, OH 43210, USA.

*To whom correspondence should be addressed. E-mail: mark.johnson@yale.edu

rangement in **C-S** are evident by cursory inspection of the overall band pattern. For example, the two free O-H stretches associated with **S** (indicated by arrows at the right of Fig. 1B) are not present in the spectrum of the **C-S** complex, indicating that both O-H groups donate H-bonds to the catalyst. The other qualitative difference is that the complex does not display the strong band near 2700 cm^{-1} clearly evident in the spectrum of **C**. That feature was tentatively assigned in an earlier survey (*20*), to an N-H^+ stretch from the ionic center engaging in an intramolecular H-bond. Its absence in **C-S** thus suggests that the internal H-bond in isolated **C** is disrupted when **S** is captured in one or more of the possible binding sites.

Beyond these cursory observations, however, it is pointless to extract detailed structures by comparing these spectra with the patterns predicted for the many calculated local minima. This is especially true as multiple isomers could be formed by the quenching procedure from room temperature to ~10 K in the ion trap. The key to our approach is to unravel which amide groups contribute to the various features in the critical amide I and amide A regions that reveal H-bond activity. This, in turn, requires that transitions can be traced to local oscillators embedded in the bimolecular assembly. At the outset, it was not obvious this would be the case as peptide spectra in solution are often considered in the context of delocalized excitons comprising contributions from multiple nearly equivalent C=O oscillators (*12*, *25*). In that case, activity in the amide I region would reflect collective motions not readily traced to a particular C=O group.

To access the degree of coupling between oscillators, we first carried out a study to monitor how site-specific placement of ^{13}C in the carbonyl of the α amide (Fig. 1) affects the suite of amide I bands. The resulting spectrum (Fig. 2) reveals that only one strong band in the amide I region (at 1640 cm^{-1}) is significantly shifted (with a weak shoulder and three minor additional perturbations much lower in vibrational frequency). This effectively isolates the location of the α-C=O stretch from overlapping bands in this region and verifies that the intrinsic width of the embedded feature is small. If the C=O stretches were intimately coupled to form a delocalized band-type structure, we would expect that removal of one of them by lowering its frequency out of the band would affect the rest. The fact that nearby bands are not affected by isotopic substitution, however, establishes that the local displacement of the α-C=O is primarily responsible for its particular transition even when the normal isotope is in place at that site. This result is not typical for the behavior of biological peptides and likely reflects the unique chemical environment created by the catalytic scaffold.

The simple response of the complex series of bands to site-specific isotopic substitution raises the prospect that many such features may be embedded in the spectra that could act as local reporters of their H-bonding environments in the spirit of the pioneering work by Rizzo and coworkers (*26*). Testing this hypothesis required synthesis of many individual, isotopically labeled analogs of the two compounds central to this study (see supporting online text for details). Figure 3 summarizes the results of site-selective isotope incorporation at four key locations within the four amide groups in **C** (see fig. S2 for raw spectra). Cursory inspection of the subtracted spectra immediately establishes that mass changes at each site contribute to only one of the amide I bands, and unambiguously identifies the specific N-H groups responsible for the two bands in the amide A region. One key feature of the particular bands that dominate the response in Fig. 3 is that they typically display shifts that approach those expected for the reduced mass change arising from substitution with ^{13}C or ^{15}N in a diatomic C=O or N-H bond (36 and 8 cm^{-1}, respectively). This highlights again that these vibrational modes are all very local oscillators, in contrast to the exciton-like, delocalized picture commonly accepted in peptide spectroscopy (*12*, *25*). The local nature of these oscillators is also confirmed by normal mode analysis in the context of harmonic spectrum calculations (see fig. S3).

The band assignments obtained from isotopic substitution are summarized with the color scheme in Fig. 3A, which includes a few contributions in the amide II region. The observed locations of the N-H donor and C=O acceptor transitions in **C** provide a particularly clear picture of the intramolecular interactions that fold the protonated peptide under the conditions of this experiment. The two horizontal arrows in trace 3A depict the range of transition frequencies corresponding to nonbonded and fully H-bonded situations for the amide I and amide A regions. The experimental spectrum is consistent with the **C** structure inferred from a computational search in a previous report (*20*) and shown schematically at the top of Fig. 3. This structure features two intramolecular H-bonds indicated by the dotted lines, in one of which the β-N-H (orange) acts as an H-bond donor and is thus red-shifted by ~130 cm^{-1} from the free δ-N-H group (green) (*27*, *28*). The third N-H is associated with the charge center involved in the intramolecular H-bond and appears near 2700 cm^{-1} (N-H^+ in Fig. 1A). We can also follow the response of the two H-acceptors, where the carbonyl groups [α (red) and γ (blue)] are found ~85 cm^{-1} lower than the two free carbonyls [δ (green) and β (orange)]. In this case, the groups engaged in H-bonding are conveniently red-shifted from those that are nonbonded and appear in otherwise clear regions of the spectrum. Consequently, the lower-energy amide I doublet is traced to the two C=O groups accepting H-bonds whereas two complementary N-H stretches (amide A) are red-shifted relative to the single nonbonded N-H group near 3500 cm^{-1}.

Having considered the band shifts associated with the intramolecular H-bonds that clamp the isolated peptide, it is a useful exercise to exploit the set of site-labeled variants of the catalyst to follow the structural changes when the excess proton is replaced by a positively charged target species with a different interaction profile. We can then follow how the various bands evolve as the catalyst binds a simpler species than the multidentate biaryl guest central to this work. The

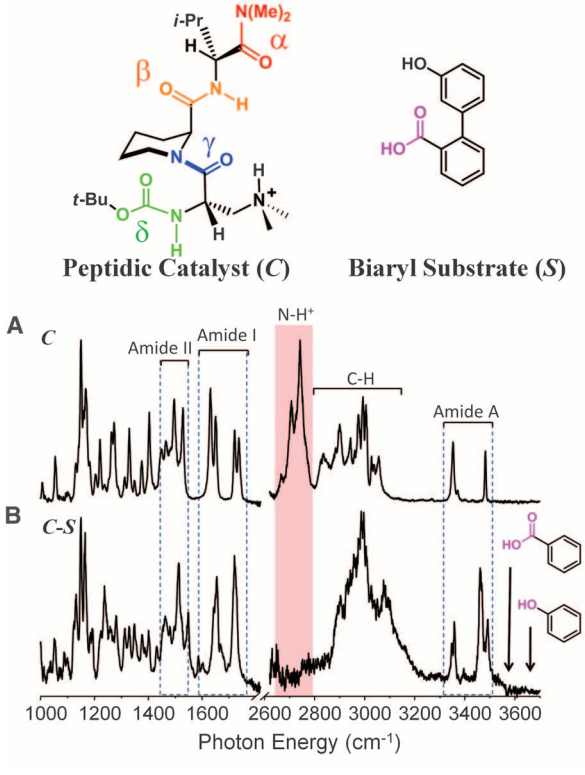

Fig. 1. The schematic structure of the peptide catalyst (**C**) and the biaryl substrate (**S**). Comparison of the infrared D_2-predissociation spectra of (**A**) **C** and (**B**) **C-S** complex. The characteristic amide regions discussed in the text are highlighted, along with the location of the intramolecular N-H^+ bond in **C**. The frequencies of free OH stretches are indicated by arrows in trace (B).

sodium ion, Na^+, provides an excellent candidate for this purpose as we expect it to strongly favor coordination to one or more of the carbonyls (*29*, *30*), and the deprotonated catalyst cannot engage in an ionic, intramolecular H-bond (largely responsible for its cyclic motif).

The D_2 predissociation spectrum of the tripeptide complexed with Na^+, $\mathbf{C}_{Na}$, is compared with that of **C** in fig. S4. The most pronounced qualitative difference is the disappearance of the intense 2720 cm^{-1} feature in the $\mathbf{C}_{Na}$ spectrum, confirming its assignment to the N-H^+ intramolecular H-bond as suggested above. The higher-frequency amide A bands are similar in the two systems, however, indicating that one of the two remaining N-H groups again participates in an intramolecular H-bond. The amide I region of $\mathbf{C}_{Na}$ still displays four distinct peaks, but their positions and intensities differ from those in **C** and the entire suite appears more compact and generally centered between the open doublets found in the **C** spectrum.

To unravel which constituents within the **C** groups are involved in the internal H-bond, as well as to establish those coordinated most closely to the ion, we followed the evolution of the various bands by using the isotopic shifts of the four sites that were studied in the protonated **C** spectra (Fig. 3). The raw data from this study are shown in fig. S5. Figure 4, A and B, compares the resulting positions of the various groups in the **C** and $\mathbf{C}_{Na}$ systems, using the same color scheme as before. Surprisingly, the most notable change going from **C** to $\mathbf{C}_{Na}$ occurs in the amide A region where the carriers of the two N-H stretch peaks are exchanged. This change implies a profound rearrangement in which the carbonyl binding of Na^+ leads the N-H of the δ amide (green) to engage in an internal H-bond while the β-N-H (orange), which acts as a donor in **C**, becomes free. Both members of the lower-energy doublet in the amide A H-bonding region are unambiguously traced to the δ amide group.

The evolution of the amide I bands from **C** to $\mathbf{C}_{Na}$ reveals the response of the C=O stretches, which are again color-coded in Fig. 4B according to amide group as indicated on the left. The α band (red), which accepted the strong ionic H-bond from the N-H^+ group in **C**, blue-shifts by 33 cm^{-1} in $\mathbf{C}_{Na}$ and changes places in energy ordering with the γ band (blue). The doublet structure of the δ amide I contribution (green) is also interesting in light of the doubling of its amide A transition, again suggesting that two closely related conformers are in play. We have confirmed that the two members of the doublet (labeled δ and δ*) are indeed due to different species by using a two-laser, isomer-selective measurement relying on photochemical hole burning (*31*), with the results shown in fig. S6. Cis and trans orientations of the *tert*-butyl (*t*-Bu) group are plausible assignments for these conformers on the basis of the calculated energies and harmonic spectra. The δ-C=O band is red-shifted relative to its position in **C** such that it overlaps the β amide I feature (orange). This points to a scenario in which three of the C=O groups interact with the ion to an appreciable extent. These qualitative features of the binding motif provide crucial constraints that narrow the computational search space and make it possible to quickly identify the local minimum that is responsible for the observed spectral features. The lowest-energy form thus obtained is displayed schematically at the left of Fig. 4B (see fig. S7 for the calculated harmonic spectra). Overall, we see that with the loss of the charged intramolecular H-bond in **C**, the peptide rearranges its conformation to accommodate the Na^+ atom by optimizing the more

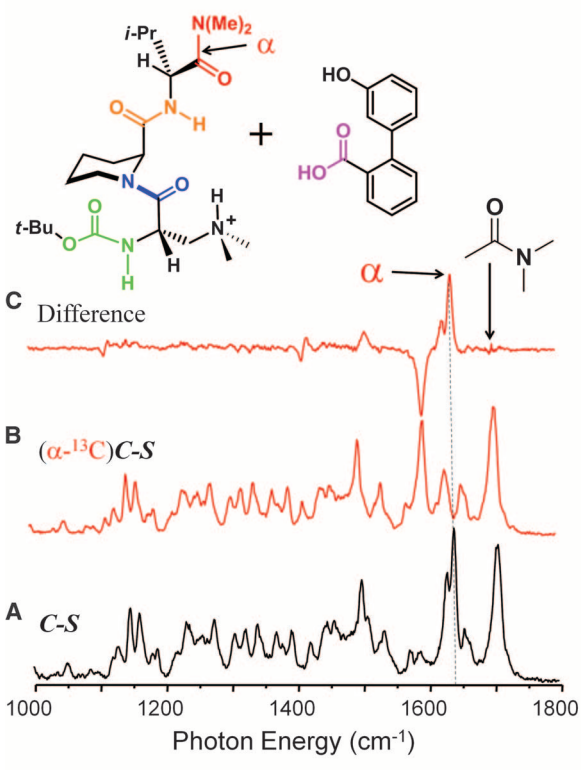

Fig. 2. IR D_2-predissociation spectrum of **C-S** with (**A**) all ^{12}C and ^{14}N, and (**B**) with α-^{13}C. The spectral changes upon isotopic labeling are highlighted in trace (**C**) which was generated by subtracting trace (B) from trace (A). Positive-going peaks in the difference spectrum indicate bands that disappear upon ^{13}C incorporation, whereas negative-going peaks reveal the location of the displaced transition, so that the mass effect manifests as a derivative-type line shape when the shift is small. The amide I frequency of isolated dimethylacetamide is indicated with an arrow.

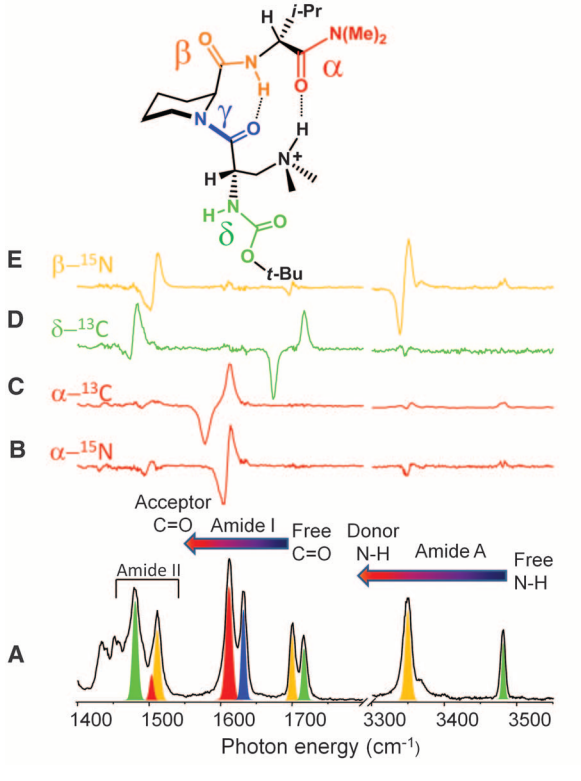

Fig. 3. (**A**) IR D_2-predissociation spectrum of **C** with all ^{12}C and ^{14}N. The spectra with heavy isotopes at the indicated positions were subtracted from trace (A) to obtain the color-coded difference spectra (**B**) to (**E**) in the four upper traces (see raw spectra in fig. S2). The color coding of the traces and peaks corresponds to the functional groups as indicated in the top schematic structure.

physical (electrostatic) interactions with three of the carbonyls (*29*, *30*). The neutral H-bond evident in the δ amide A region is then traced to the nearby tertiary amine group, while the β-N-H is rotated free as its C=O group pivots away from the ion.

We now turn to the docking motif that tethers the biaryl substrate to the peptide scaffold in the noncovalently bound complex, **C-S**. The absence of free O-H transitions (Fig. 1B) already established that both O-H groups on **S** are engaged in donor H-bonds. The positions of the **S**–CO_2H–based transitions are established by following the effect of ^{13}C placement in the acid group, with the difference spectrum displayed in fig. S8. This spectrum reveals that the acid C=O stretch is a dominant component of the higher-energy amide I feature at 1704 cm^{-1} (highlighted in purple in the color-coded spectrum in Fig. 4C). This carboxylic acid–derived transition falls about 50 cm^{-1} below the 1752 cm^{-1} value reported (*32*, *33*) for isolated benzoic acid (labeled in Fig. 4C), indicating that its C=O component is engaged as an acceptor of either a charged single H-bond or multiple neutral H-bonds. There are, indeed, two closely spaced bands in the neutral H-bonded N-H region near 3350 cm^{-1} in Fig. 4C that could in principle be engaged in such a scenario. However, the site-specific difference spectra unambiguously trace both members of this doublet to the same β-N-H group, and thus, like the $\mathbf{C}_{Na}$ system, this doubling likely results again from two isomeric forms of the adduct (labeled β and β*) in Fig. 4C. The other NH group must therefore reside in a nonbonding site; we can rule out the double donor possibility and conclude that the acid carbonyl binds to the protonated amine.

Because the basicity of the substrate is much smaller than that of the deprotonated catalyst, the complexation cannot dislodge the excess proton from its point of attachment at the tertiary amine in **C**. Although ionic H-bonding to the acid group may appear counterintuitive at first glance, it is in fact typical for organic acids, RCO_2H, to share an excess proton between their carbonyl groups in the proton-bound binary complexes ($RCOH{=}O\cdot\cdot H^+\cdot\cdot O{=}COHR'$) (*34*, *35*). This propensity is consistent with the preferential protonation of the bare carboxylic acids at the C=O to form diol-type arrangements [$R\text{-}(OH)_2^+$]. When the α-C=O is displaced by –CO_2H on **S**, the N-H^+ band is expected to blue-shift due to the lower basicity of the acid group relative to that of the (α) amide (790.1 versus 877.0 kJ/mol for gas-phase basicity, respectively) (*36*). The nature of the acid functionality is further defined by the behavior of the weaker bands at lower energy, which probe the motion of the C-OH moiety. The difference spectrum in fig. S8 reveals a transition affected by ^{13}C substitution in the acid group near 1250 cm^{-1}. This wide splitting of the two bands associated with –CO_2H (1700 and 1250 cm^{-1}) confirms that the general structure of the acid is intact (i.e., HO-C=O as opposed to a salt-bridge motif involving –CO_2^-) and is therefore acting as a strongly H-bonded domain that both accepts and donates (recall the missing free O-H discussed above) an H-bond to **C**.

To address the roles of the H-bonding sites on **C**, we again followed the evolution of the specific groups using the four isotope labels as before (raw data are in fig. S9). Like the case of the α amide depicted in Fig. 1, all sites yield definitive assignments of bands associated with the N-H donors and C=O acceptors, with the results included in the color-coded segments of the amide I and A regions in Fig. 4C. The α-C=O in **C-S** is blue-shifted relative to its position in **C**, consistent with the loss of the ionic H-bond to N-H^+, but still lies well below the range for a free C=O stretch. As a result, we conclude that it accepts a neutral H-bond. The remaining labeled transitions on the catalyst establish that the β-N-H acts as a donor whereas the δ-N-H is free. The δ-C=O (green) is red-shifted (by 20 cm^{-1}) from its uncomplexed position, establishing that it is an H-bond acceptor. Application of these constraints markedly reduces the number of possible **C-S** structures, which serves to efficiently guide the ab initio search for minimum-energy structures consistent with the empirically deduced H-bonding pattern. The most important clue derives from the assignment of the attachment occurring between the protonated amine on **C** and the acid carbonyl on **S**. Subsequent locking of the two relatively floppy moieties linked by the shared proton then involves only a relatively small number of plausible arrangements anticipated by force-field models.

These candidate structures are then sorted by the pattern of H-bonding, which reveals the only **C-S** structure consistent with the active groups presented in Fig. 4D. About 20 low-lying isomers were recovered by using electronic structure theory to identify the locally stable conformations available to the system. The structure in Fig. 4D is the global minimum identified in this computational search, being lower in energy by 0.2 eV (19.3 kJ/mol) relative to the next lower structure. The harmonic spectrum associated with this structure is indicated in Fig. 4D, which is in excellent agreement with the bands assigned by isotopic labeling. Hence, we propose the plausible assignments of other features not directly pinned down by empirical behavior of the isotopomers by the color coding in Fig. 4C. These predictions provide a solid foundation with which to confirm the structure by further labeling if desired.

The most important feature of the **C-S** structure shown in Fig. 4, C and D, is that, although different from the salt-bridge arrangement suggested to explain stereoselective bromination (*2*), the gas-phase ion complex indeed places the phenol constituent in close proximity to the catalytic dimethyl formamide functionality. Also, this docking motif is quite robust, as all site-selective

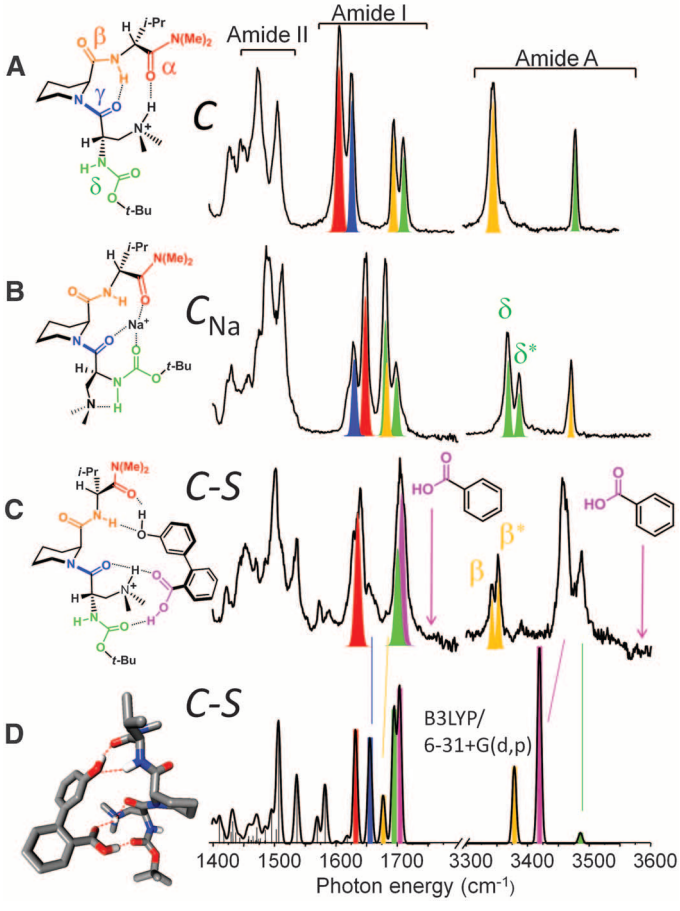

Fig. 4. Summary of amide I and A peak assignments in the **(A) C**, **(B) $\mathbf{C}_{Na}$**, and **(C) C-S** IR spectra. The color coding of the peaks corresponds to the functional groups highlighted in the corresponding schematic structures on the left. Trace **(D)** presents the **C-S** harmonic spectrum calculated at the B3LYP/6-31+G(d,p) level. The **C-S** lowest-energy structure is displayed at the left of trace (D) and displayed schematically at the left of trace (C) to highlight the interactions.

difference spectra are quite simple, with only one main band (or very close doublet) affected by substitution. Consequently, only one dominant configuration (with minor conformers arising from remote group orientation) is adopted upon extraction from solution. It is possible that removal of solvent changes the structure, and that the active species in solution could be electrically neutral, and both of these issues can be addressed with the cryogenic approach by retaining many solvent molecules and working with bulky noncoordinating counterions such as tetramethyl ammonium to explore the neutral adduct. It is nonetheless an important first step to establish that in the protonated adduct characterized here, the substrate (**S**) adopts a configuration consistent with the observed enantioselectivity of the catalyst (*2*, *5*).

The pivotal assignment of the strong contact occurring between N-H^+ and CO_2H in **C-S** relies solely on the behavior of the ^{13}C label in the acid position. This raises the importance of not only identifying whether particular groups are involved in H-bonding, but also establishing specifically which donors and acceptors are paired, at least in one contact point. One avenue to explore, therefore, is whether the intensities of both the donor and acceptor bonds can be modulated by the isotopic labeling scheme. An attractive possibility in this regard would be to monitor the combination band involving one quantum in both the donor and acceptor groups (e.g., N-H and C=O), which one would expect to be activated by the anharmonic coupling inherent in the H-bond. The presence of an isotope shift from both ^{15}N and ^{13}C labels would then confirm a specific point contact.

This procedure yields a microscopic picture of a docking arrangement, sufficiently constraining the possible structures such that electronic structure theory can be efficiently used to converge on a unique minimum-energy structure within an otherwise computationally prohibitive, high-dimensionality landscape. The method appears general and likely to become a central tool for the characterization of processes that depend on supramolecular associations.

References and Notes

1. D. H. Williams, B. Bardsley, *Angew. Chem. Int. Ed.* **38**, 1173 (1999).
2. J. L. Gustafson, D. Lim, S. J. Miller, *Science* **328**, 1251 (2010).
3. M. S. Taylor, E. N. Jacobsen, *Angew. Chem. Int. Ed.* **45**, 1520 (2006).
4. A. G. Doyle, E. N. Jacobsen, *Chem. Rev.* **107**, 5713 (2007).
5. J. L. Gustafson, D. Lim, K. T. Barrett, S. J. Miller, *Angew. Chem. Int. Ed.* **50**, 5125 (2011).
6. J. M. Lehn, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4763 (2002).
7. K. H. Bleicher, H. J. Böhm, K. Müller, A. I. Alanine, *Nat. Rev. Drug Discov.* **2**, 369 (2003).
8. J. Bajorath, *Nat. Rev. Drug Discov.* **1**, 882 (2002).
9. W. L. Jorgensen, *Science* **303**, 1813 (2004).
10. B. K. Shoichet, *Nature* **432**, 862 (2004).
11. L. J. Prins, D. N. Reinhoudt, P. Timmerman, *Angew. Chem. Int. Ed.* **40**, 2382 (2001).
12. S. Ham, S. Cha, J. H. Choi, M. Cho, *J. Chem. Phys.* **119**, 1451 (2003).
13. C. L. Perrin, J. B. Nielson, *Annu. Rev. Phys. Chem.* **48**, 511 (1997).
14. G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer-Verlag, Berlin, New York, 1991).
15. P. I. Haris, G. T. Robillard, A. A. van Dijk, D. Chapman, *Biochemistry* **31**, 6279 (1992).
16. S. Sonar *et al.*, *Nat. Struct. Biol.* **1**, 512 (1994).
17. L. Tadesse, R. Nazarbaghi, L. Walters, *J. Am. Chem. Soc.* **113**, 7036 (1991).
18. R. R. Knowles, E. N. Jacobsen, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20678 (2010).
19. M. Z. Kamrath, R. A. Relph, T. L. Guasco, C. M. Leavitt, M. A. Johnson, *Int. J. Mass Spectrom.* **300**, 91 (2011).
20. M. Z. Kamrath *et al.*, *J. Am. Chem. Soc.* **133**, 6440 (2011).
21. M. Okumura, L. I. Yeh, J. D. Myers, Y. T. Lee, *J. Chem. Phys.* **85**, 2328 (1986).
22. D. J. Goebbert, T. Wende, R. Bergmann, G. Meijer, K. R. Asmis, *J. Phys. Chem. A* **113**, 5874 (2009).
23. T. R. Rizzo, J. A. Stearns, O. V. Boyarkin, *Int. Rev. Phys. Chem.* **28**, 481 (2009).
24. J. A. Stearns *et al.*, *J. Am. Chem. Soc.* **129**, 11814 (2007).
25. L. Wang *et al.*, *J. Am. Chem. Soc.* **133**, 16062 (2011).
26. J. A. Stearns, C. Seaiby, O. V. Boyarkin, T. R. Rizzo, *Phys. Chem. Chem. Phys.* **11**, 125 (2009).
27. R. R. Gardner, G. B. Liang, S. H. Gellman, *J. Am. Chem. Soc.* **117**, 3280 (1995).
28. C. P. Rao, R. Nagaraj, C. N. R. Rao, P. Balaram, *Biochemistry* **19**, 425 (1980).
29. D. Semrouni *et al.*, *J. Am. Soc. Mass Spectrom.* **21**, 728 (2010).
30. R. C. Dunbar, J. D. Steill, J. Oomens, *Int. J. Mass Spectrom.* **297**, 107 (2010).
31. B. M. Elliott *et al.*, *J. Chem. Phys.* **129**, 094303 (2008).
32. J. M. Bakker, L. Mac Aleese, G. von Helden, G. Meijer, *J. Chem. Phys.* **119**, 11180 (2003).
33. S. G. Stepanian, I. D. Reva, E. D. Radchenko, G. G. Sheina, *Vib. Spectrosc.* **11**, 123 (1996).
34. Y. Inokuchi, N. Nishi, *J. Phys. Chem. A* **107**, 11319 (2003).
35. H. M. Lee *et al.*, *Phys. Chem. Chem. Phys.* **12**, 6278 (2010).
36. E. P. L. Hunter, S. G. Lias, *J. Phys. Chem. Ref. Data* **27**, 413 (1998).

Acknowledgments: M.A.J. thanks the Air Force Office of Scientific Research under grant FA-9550-09-1-0139. We also thank the U.S. NSF under grants CHE-091199 (M.A.J.) and CHE-0848242 (A.B.M.). S.J.M thanks the NIH under grant R01-GM068649. This work was supported in part by the Yale University Faculty of Arts and Sciences High Performance Computing facility (and staff).

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1214948/DC1
Materials and Methods
Figs. S1 to S9
References (*37*–*40*)

7 October 2011; accepted 20 December 2011
Published online 19 January 2012;
10.1126/science.1214948

A Molecular MoS_2 Edge Site Mimic for Catalytic Hydrogen Generation

Hemamala I. Karunadasa,[1,2] Elizabeth Montalvo,[1] Yujie Sun,[1,2] Marcin Majda,[1] Jeffrey R. Long,[1,3]* Christopher J. Chang[1,2,4]*

Inorganic solids are an important class of catalysts that often derive their activity from sparse active sites that are structurally distinct from the inactive bulk. Rationally optimizing activity is therefore beholden to the challenges in studying these active sites in molecular detail. Here, we report a molecule that mimics the structure of the proposed triangular active edge site fragments of molybdenum disulfide (MoS_2), a widely used industrial catalyst that has shown promise as a low-cost alternative to platinum for electrocatalytic hydrogen production. By leveraging the robust coordination environment of a pentapyridyl ligand, we synthesized and structurally characterized a well-defined Mo^{IV}-disulfide complex that, upon electrochemical reduction, can catalytically generate hydrogen from acidic organic media as well as from acidic water.

Molybdenite, or molybdenum disulfide (MoS_2), the earliest form of molybdenum to be identified from ores, is one of the most widely used catalysts in industry today as the standard for hydrodesulfurization (HDS) of petroleum (*1*). In its nanoparticulate form, MoS_2 has further demonstrated promise as an inexpensive alternative to platinum for the electrochemical and photochemical generation of hydrogen from water (*2*–*6*). As is the case with many inorganic solids, the catalytic activity of MoS_2 is localized to rare surface sites, whereas the bulk material is relatively inert (*7*–*9*). High-resolution scanning tunneling microscopy studies and theoretical calculations performed on nanoparticulate MoS_2 structures that form under sulfiding conditions implicate the formation of disulfide linkages or triangular MoS_2 units along the fully sulfided catalytically active edges of the layered structure (*10*–*14*). However, the precise molecular structures and modes of action of these sites remain elusive. Because of the bulk material's layered structure, which favors the growth of plate-like crystals, a single crystal with a large edge dimension is extremely challenging to prepare (*1*). Here, we report the synthesis of a well-defined molecular analog of the proposed MoS_2 edge structure, a side-on bound Mo^{IV}-disulfide complex. Electrochemical reduction of this mol-

[1]Department of Chemistry, University of California, Berkeley, CA 94720, USA. [2]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [3]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [4]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: chrischang@berkeley.edu (C.J.C.); jrlong@berkeley.edu (J.R.L.)