REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-04-2008	Final Technical Report	14/07/06-31/01/08

4. TITLE AND SUBTITLE
Bio-Inspired Materials: Protein Cage Architectures

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N0014-06-1-1016

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Douglas, Trevor

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Montana State University
113 Chemistry and Biochemistry Building
Bozeman, MT 59717

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Montana State University
Office of Sponsored Program
309 Montana Hall
Bozeman, MT 59717

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

20080501236

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Trevor Douglas
					19b. TELEPHONE NUMBER *(Include area code)* 406-994-6566

DEFENSE TECHNICAL INFORMATION CENTER

Information for the Defense Community

DTIC® has determined on Month 05 Day 08 Year 2008 *that this Technical Document has the Distribution Statement checked below. The current distribution for this document can be found in the DTIC® Technical Report Database.*

☒ **DISTRIBUTION STATEMENT A.** Approved for public release; distribution is unlimited.

☐ © **COPYRIGHTED.** U.S. Government or Federal Rights License. All other rights and uses except those permitted by copyright law are reserved by the copyright owner.

☐ **DISTRIBUTION STATEMENT B.** Distribution authorized to U.S. Government agencies only. Other requests for this document shall be referred to controlling office.

☐ **DISTRIBUTION STATEMENT C.** Distribution authorized to U.S. Government Agencies and their contractors. Other requests for this document shall be referred to controlling office.

☐ **DISTRIBUTION STATEMENT D.** Distribution authorized to the Department of Defense and U.S. DoD contractors only. Other requests shall be referred to controlling office.

☐ **DISTRIBUTION STATEMENT E.** Distribution authorized to DoD Components only. Other requests shall be referred to controlling office.

☐ **DISTRIBUTION STATEMENT F.** Further dissemination only as directed by controlling office or higher DoD authority.

Distribution Statement F is also used when a document does not contain a distribution statement and no distribution statement can be determined.

☐ **DISTRIBUTION STATEMENT X.** Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoDD 5230.25.

Format for Final Technical Report, ONR Code 341 (Biology)

Please limit the total report to 6 pages. Please use 12 point font with 1' margins on all edges. ATTACH ONE COPY OF ALL CITED PAPERS AND ABSTRACTS NOT PREVIOUSLY SUBMITTED TO ONR.

Cover Page:

GRANT # : N00014-06-1-1016

PRINCIPAL INVESTIGATOR : Dr. Trevor Douglas

INSTITUTION : Montana State University

GRANT TITLE : Bio-Inspired Nanomaterials: Protein Cage Nano-Architectures

AWARD PERIOD : 15 June 2006 – 31 January 2008

OBJECTIVE :

The vision for the Center for Bio-Inspired Nanomaterials (CBIN) at Montana State University builds upon a strong scientific foundation and a proven multidisciplinary team in the development of protein cage architectures as templates for nanomaterials synthesis and applications in nanotechnology. The exterior interface, the interior interface, and the interface between protein subunits can all be manipulated in the pursuit of refined nanoparticle properties. The synthesis and characterization capabilites of CBIN will be developed as a foundation to provide nanomaterials with novel optical, electronic, magnetic, and acoustic properties. These materials will be investigated as the basis for applications in magnetic sensing, biofilm antifouling, and nanocatalysis. Our previous success in utilizing protein cages as size and shape constrained templates for synthesis of a wide range of nanostructures has led to an understanding that protein cage dynamics plays an important role in the overall properties of the materials. The long-term goal is to use this knowledge to guide the development of a new generation of active and responsive nanomaterials. The directed program objectives of CBIN are:

- *To use protein cages and architectures for the formation of nano-materials with applications in magnetic sensing.*
- *To use protein cages and architectures to create nano-catalysts with high specificity and efficiency for hydrogen production and metal ion remediation.*
- *To use the library of protein cages and architectures to create nanomaterials for targeting biofilms for imaging and anti-fouling.*

APPROACH: The Center for Bio-Inspired Nanomaterials (CBIN) has facilitated the development of a campus-wide infrastructure for the training of personell (post-docs, graduate students, and undergraduates) and for collaborative multidisciplinary research. Through CBIN, the infrastructure is in place for large-scale production and synthesis of materials and their characterization. Materials are produced using genetic and chemical modification of protein cage materials and controlled chemical synthesis under mild biological conditions. High-resolution structural techniques (TEM, X-ray crystallography, NMR) and spectroscopy to probe electronic structure (XAS, EPR, ACMS) will be coupled with functional investigation of the materials properties. Lithography

ACCOMPLISHMENTS:

Protein Cage Architectures for Applications in Nanomaterials

Protein cage architectures provide a unique platform for synthesis of size and composition constrained nanomaterials. We have shown that we can utilize charged interfaces in the protei to direct nucleation and crystallization of nanomaterials constrained by a variety of protein cage architectures. In particular, using this approach we have shown the size constrained synthesis of magnetic materials by growing nanoparticles of transition metal oxide materials. These materials show unusual size dependent properties. Recently we have demonstrated that the multifunctional protein cage platform can be used to attach individual paramagnetic centers in a site specific manner to a viral protein cage. The magnetic behaviour of thes resulting materials

show some remarkable and unexpected behaviour. Firstly using a genetically incorporated peptide for Gd binding we were able to make a viral capsid showing very high proton relaxivity. In contrast, the attachment of metal chelates to the viral capsid (at very high density – 180 per cage) showed significantly lower proton relaxivity. Our analysis shows that local mobility of the Gd ion plays a huge role in the composite magnetic properties of the capsid-based material. Future work is directed at utilizing this observation/analysis to make a new class of materials with directed magnetic properties based on a combination of controlled mobility and metal ligand interactions.

Using the exterior surface of the CCMV viral cage we have chemically attached ligands that direct the cage specifically to the surface of a bacterial cell. The targeting moiety used in this case was an antibody specific to protein A expressed on the surface of the biofilm forming bacteria, Stapholococcus aureus. As part of our efforts to produce active multifunctional nanomaterials a photosensitizor , a $(Ru(bpy)_3^{2+}$ derivative. This photosensitizor acts as a catalyst for singlet oxygen production and results in the selective killing of S. aureus only when coupled to the cell specific targeting. The importance of these results lies in our demonstrated ability to combine self-assembly, targeting, and functional catalytic activity into a single nanoparticle platform. Future work will be directed towards targeting of other DoD relavant biofilm forming organisms together with delivery of selectively antibacterial agents.

Protein Cage Dynamics

Protein function is intimately connected to dynamics and therefore knowledge of the frequency, range, and coordination of motion by supramolecular complexes is critical to understanding function and the development of bio-inspired nanomaterials. We are using protein cages as a paradigm for studying protein dynamics in supramolecular complexes. The extremely large size and icosahedral architecture of virus capsids limit the use of many standard techniques for studying protein motion. To overcome these problems, we employ an array of biophysical techniques to investigate the solution phase behavior of viruses. Kinetic hydrolysis, an approach being developed in the Bothner lab, is a straight-forward and powerful technique for identifying the dynamic regions within a single protein or in the context of a multi-component complex. Protein dynamics is being investigated at three levels: the dynamics of the subunit, the assembled cage architecture, and the dynamics associated with higher order particle/particle and surface/particle interactions.

The function of viral capsids, and macromolecular complexes in general, requires conformational freedom. The work that has been accomplished in the past year contains a number of major technical advances that allow us to understand the assembly and function of a protein cage, Hepatitis B Virus.
These include:

1. The first rigorous measurements of localized rates and equilibria for dynamic motion within a megadalton complex.
2. The biologically important C-terminal region of HBV capsid protein is dynamic in both dimer and capsid forms. This is in contrast to the assignment as a well-ordered helix in X-ray and cryo-EM based models.

3. This region is a switch that regulates capsid assembly and function, with thermodynamic profiles that invert depending on the assembly state.
4. Enzymatic hydrolysis (kinetic hydrolysis), which has to date been applied largely to model systems, can be extended to study large multi-component complexes: this biophysical technique may be the most relevant tool for studying the structural and stability of nanomaterials.

Hydrogen Catalyst Thrust:

Nature has evolved very active catalysts for H_2 production in the form of the Fe-containing hydrogenases which can produce 6000-9000 molecules of hydrogen per second for a single molecule of enzyme. Light energy can be harnessed and coupled to the reduction of a hydrogenase using an electron transfer mediator such as methyl viologen using light activated porphyrins or ruthenium (II) complexes. We have been able to achieve hydrogen production in an in-vitro assay containing hydrogenase, semiconductors (or molecular light harvesting antennae) with methyl viologen as a redox mediator and simple organic acids as a primary source of reducing equivalents. The in-vitro complex system serves as an ideal platform for the characterization and optimization of hydrogen production catalysts, photocatalysts, and coupling mechanisms. We have been examining in detail the structural features of examples of members of both the [NiFe]- and [FeFe]- classes of hydrogenase. These studies are aimed at 1) understanding the structural determinants for catalysis which provide critical information for the design and synthesis of biomimetic catalysts and 2) understanding the structural determinants of enzyme stability providing the necessary information for the identification and engineering of stable hydrogenases to be utilized in biohybrid materials.

1) Our [FeFe]-hydrogenase structural studies are at an advanced stage with a number of published structures of the enzyme from *Clostridium pasteurianum* are currently working the detailed structural analysis at better than 1.4 Å resolution. The preliminary analysis of this structure together with complementary computational studies conducted by the Szilagyi group have provide the basis for critical examination of the composition of the unique dithiolate ligand of the cluster with paradigm challenging implications. The suggestion that the cluster exists with an ether linked dithiolate has significant impact on mechanistic considerations of H cluster catalyzed reversible hydrogen oxidation and provides the basis for the design of mimetic and biochemical experiments to examine ligand assignments directly.

2) Another avenue of our work has focused on using the high resolution structure as well as genomic analysis of the variances in amino acid sequence of hydrogenases from different microbial sources as a platform for identifying the structural determinants for hydrogenase stability and identifying appropriate model systems for use in materials synthesis. We have compared through homology modeling, [FeFe]-hydrogenases from both mesophilic and thermophilic as well as from strict and facultative anaerobes to develop hypotheses for enzyme thermal adaptation and oxygen tolerance. In addition, we have worked from the observations made by Posewitz, King, and coworkers in which it was shown that [FeFe]-hydrogenases can be successfully expressed in *E. coli* when the structural gene products for the enzyme are expressed in a background in which the accessory enzymes HydE, HydG, and HydF are coexpressed. We

have been using this expression system to examine the heterologous expression of active hydrogenases from a variety of microbial sources. To date, we have examined [FeFe]-hydrogenases from five different thermophile sources. We have constructed the appropriate clones to heterologously express a number of hydrogenases from thermophiles and hyperthermophiles. Unfortunately, to date we have been unsuccessful in the heterologous expression of [FeFe]-hydrogenases with detectable [FeFe]-hydrogenases, an interesting observation since the system has worked well for a number of hydrogenases from mesophiles. These observations have motivated us to revisit the amino acid sequence and genomic context of thermophile [FeFe]-hydrognease as well as the overall metabolism of thermophiles possessing these [FeFe]-hydrogenases. The results of this analysis suggests that the [FeFe]-hydrogenases in these organisms may be functioning physiologically as hydrogen oxidation catalysis in these organisms. In addition, amino acid sequence differences are noted at key positions in the enzymes active sites of thermophile [FeFe]-hydrogenases when compared to their mesophile counterparts. We are currently moving forward with the working hypothesis that the subtle differences allow the [FeFe]-hydrogenases to function more effectively in hydrogen oxidation in these thermophile hosts and we are testing this hypothesis by amino acid substitution experiments.

Magnetic Nanoparticle thrust:

The first major accomplishments for this period, reported in the Journal of Applied Physics, is the extension of our work in the application of the first principles understanding of magnetic damping for itinerant ferromagnets. We have shown that there is a strong correlation between the density of states and the damping rate and have shown that two mechanisms are responsible for the temperature dependence of the damping, one an inter-band and the second an intra-band excitation. We can now calculate, from first principles the damping rate as the electronic bandstructure is modified through doping or strain, allowing for a predictive tool for engineered damping structures.

A second major accomplishment, submitted to the Journal of Magnetism and Magnetic Materials, is our development of angle dependent Electron Magnetic Resonance measurements of different sized Fe-oxide nanoparticles synthesized inside different sized protein cage structures. By cooling a suspension of encapsulated magnetic nanoparticles in an applied field below the solution freezing temperature, we are able to show that the nanoparticles are locked to the protein cage structure and that the surface anisotropy term is large. As the particle size is decreased, the increase in the surface-to-volume ratio allows us to separate the contribution of the surface anisotropy from the bulk anisotropy. This has important implications in using magnetic nanoparticles as elements of a sensor technology.

A third major accomplishment, not submitted yet, is the identification of the doping site for Mn and Zn doping (in a range of 0-33%) of Fe-oxide magnetic nanoparticles. By using X-ray absorption spectroscopy and magnetic circular dichroism, we find that both the Zn and the Mn substitute into an octahedral site with a +2 valance state. Because of this substitution, the magnetic moment is controllable from 32 m_B per unit cell down to zero. This demonstrates that the magnetic moment of these particles, as well as the anisotropy values, are controllable through doping.

We have developed new methods to characterize the cage structures and the mineralized cores, both while encapsulated and after protein removal, for the successful synthesis of uniform magnetic nanoparticles of controllable shape, size, composition, and magnetic properties. We have successfully generated microscale patterns of antibody and protein cage on silicon and gold substrates using both microplotter and FIB-based patterning technologies. The results are as follows:

Patterning by microplotter was achieved by depositing a preselected antibody solution directly onto chemically activated silicon or gold surfaces after modifying these surfaces with *maleimido-* or *NHS-* groups. The spot size achieved by micro-plotter is a function of many parameters, including the tip diameter of the micro-plotter, the surface hydrophobicity, and the optimization of experimental parameters such as the quality of surface modification, concentration of proteins, plotting speed, and voltage applied to the driving piezo. Using an optimized microplotter, microarrays of bacterial cells with a spot size of 25 μm have been achieved on *maleimido-* modified silicon surfaces, an example of which is shown in Fig. 1a below.

FIB-based patterning was achieved using a focused Ga+ beam to create patterns on a silicon or gold substrate that was pre-passivated with molecules that resist the nonspecific binding of proteins. We have successfully achieved a lateral resolution of ~1 μm and can prepare patterns with sizes anywhere from 1 μm to 200 μm. We applied this approach to the immobilization of *Sulfolobus* turreted icosahedral virus (STIV), a cage-shaped protein with a diameter of ~ 50nm, on pre-modified silicon substrates. Although the attachment of STIV cages to silicon was observed, we observed low coverage density of STIV on the surface, presumably due to the poor purity of the STIV cages used. We believe this is because the STIV sample used in our experiments most likely contained not only the protein cages but also other small proteins that competed with the attachment of STIV to the substrate surface. This competitive binding reduced the density of STIV cages on the surface. We expect to improve the density by using high purity protein cages (i.e. tobacco mosaic virus (TMV). This work is under active consideration as time permitting.

Computational Thrust:

A systematic data collection was carried out at the Stanford Synchrotron Radiation Laboratory for commercially available, freshly prepared bulk, and protein-cage assisted synthetic iron and molybdenum oxide and sulfide, samples. Qualitative data comparison further support earlier observations by CBIN members that protein cages allow for accessing areas of iron oxide/sulfide phase diagrams that would not be possible using conventional synthetic methods. Using the complementarity of the X-ray absorption data collected at the iron K-edge (EXAFS), iron L-, and ligand K-edges (XANES), we began quantitatively determine the structural differences among various species.

With respect our computational work we joined efforts with the Peters group on mapping the soft/hard material interface of iron-sulfur clusters embedded into a protein matrix. In this collaboration, we employed computational chemical tools and approaches developed by us to probe a key question related to the active site of biological hydrogen production. We carried out an unbiased investigation of the chemical composition of the catalytically active iron-sulfur

cluster of FeFe-hydrogenase and found indications for a new chemical composition that has not been critically explored experimentally. The CBIN funding was essential to achieve this allowing the establishment of a virtual chemical model of the active site that captures the most important interactions between the iron-sulfur active site and protein environment.

SIGNIFICANCE : CBIN has facilitated the collaborative interaction between biologists, biochemists, chemists, and physicists focused on the use and understanding of constrained reaction environments provided by viruses and other protein cage architectures. This bio-inspired approach utilizes and exploits hard-soft interfaces, particle-surface interactions and controlled hierarchical assembly for the synthesis of functional nanomaterials. The underlying design principles are based on an understanding of biological systems. In addition, we are also using biological materials as templates for the construction of materials with superior properties. Thus, control over material synthesis results in magnetic materials with unique properties, catalytic materials with higher efficiency, and materials that can be used to both integrate and kill bacterial biofilms, superior to those produced by other methods. We have established a productive team that spanning many disciplines and able to train the next generation of scientists and respond to the materials needs of the nation. The overall significance of this work lies in the combination of fundamental scientific understanding coupled to directed applications of national importance.

PATENT INFORMATION : A patent entitled "Novel Nanoparticles for Biofilm Targeting" Serial # 60/830,830 has been filed with inventors Drs. Trevor Douglas, Mark Young, and Peter Suci.

PUBLICATIONS and ABSTRACTS (for total period of grant):

Publications in **BOLD** have not previously been submitted to ONR and are attached.

1. Uchida, M.; Klem, M.; Allen, M.; Flenniken, M.; Gillitzer, E.; Varpness, Z.; Suci, P.; Liepold, L.; Young, M.; Douglas, T., Biological Containers: Protein Cages as Multifunctional Nanoplatforms. *Advanced Materials* 2007, 19, 1025-1042.

2. Suci, P. A.; Berglund, D. L.; Liepold, L.; Brumfield, S.; Pitts, B.; Davison, W.; Oltrogge, L.; Hoyt, K. O.; Codd, S.; Stewart, P. S.; Young, M.; Douglas, T., High-density targeting of a viral multifunctional nanoplatform to a pathogenic, biofilm-forming bacterium. *Chemistry & Biology* 2007, 14, (4), 387-398.

3. **Suci, P.; Varpness, Z.; Gillitzer, E.; Douglas, T.; Young, M., Targeting and photodynamic killing of a microbial pathogen using protein cge architectures functionalized with a photosensitizer. *Langmuir* 2007, 23, (24), 12280-12286.**

4. Negusse, E.; Lussier, A.; Dvorak, J.; Idzerda, Y. U.; Shinde, S. R.; Nagamine, Y.; Furukawa, S.; Tsunekawa, K.; Djayaprawira, D. D., Magnetic characterization of CoFeB/MgO and CoFe/MgO interfaces. *Applied Physics Letters* 2007, 90, (9).

5. **Liepold, L.; Anderson, S.; Willits, D.; Oltrogge, L.; Frank, J.; Douglas, T.; Young, M., Viral Capsids as MRI Contrast Agents. *Magnetic Resonance in Medicine* 2007, 58, (5), 871-879.**

6. Klem, M. T.; Resnick, D. A.; Gilmore, K.; Young, M.; Idzerda, Y. U.; Douglas, T., Synthetic control over magnetic moment and exchange bias in all-oxide materials encapsulated within a spherical protein cage. *Journal of the American Chemical Society* 2007, 129, (1), 197-201.

7. **Gilmore, K.; Idzerda, Y.; Stiles, M., Identification of the Dominant Precession-Damping Mechanism in Fe, Co, and Ni by First Principles Calculations. *Physical Review Letters* 2007, 99, 027204.**

8. Uchida, M.; Flenniken, M. L.; Allen, M.; Willits, D. A.; Crowley, B. E.; Brumfield, S.; Willis, A. F.; Jackiw, L.; Jutila, M.; Young, M. J.; Douglas, T., Targeting of cancer cells with ferrimagnetic ferritin cage nanoparticles. *Journal of the American Chemical Society* 2006, 128, (51), 16626-16633.

9. Suci, P. A.; Klem, M. T.; Arce, F. T.; Douglas, T.; Young, M., Assembly of multilayer films incorporating a viral protein cage architecture. *Langmuir* 2006, 22, (21), 8891-8896.

10. Rokhsana, D.; Dooley, D. M.; Szilagyi, R. K., Structure of the oxidized active site of galactose oxidase from realistic in silico models. *Journal of the American Chemical Society* 2006, 128, (49), 15550-15551.

11. Merz, M.; Roth, G.; Reutler, P.; Buchner, B.; Arena, D.; Dvorak, J.; Idzerda, Y. U.; Tokumitsu, S.; Schuppler, S., Orbital degree of freedom in single-layered La1-xSr1+xMnO4: Doping- and temperature-dependent rearrangement of orbital states. *Physical Review B* 2006, 74, (18).

12. Merz, M.; Reutler, P.; Buchner, B.; Arena, D.; Dvorak, J.; Idzerda, Y. U.; Tokumitsu, S.; Schuppler, S., O1s and Mn2p NEXAFS on single-layered La1-xSr1+xMnO4: crystal field effect versus orbital coupling mechanism. *European Physical Journal B* 2006, 51, (3), 315-319.
13. Maaty, W. S. A.; Ortmann, A. C.; Dlakic, M.; Schulstad, K.; Hilmer, J. K.; Liepold, L.; Weidenheft, B.; Khayat, R.; Douglas, T.; Young, M. J.; Bothner, B., Characterization of the archaeal thermophile Sulfolobus turreted icosahedral virus validates an evolutionary link among double-stranded DNA viruses from all domains of life. *Journal of Virology* 2006, 80, (15), 7625-7635.
14. Gillitzer, E.; Suci, P.; Young, M.; Douglas, T., Controlled ligand display on a symmetrical protein-cage architecture through mixed assembly. *Small* 2006, 2, (8-9), 962-966. mineralized in protein cages. *Journal of Applied Physics* 2005, 97, (10).
15. **Suo, Z; Avci, R; Yang, X; Pascual, D.W.; Efficient Immobilization and Patterning of Live Bacterial Cells. *Langmuir* 2008, 24, 4161-4167.**
16. Hilmer, J.K.; Zlotnick, A.; Bothner, B.; Conformational Equilibria and Rates of Localized Motion within Hepatitis B Virus Capsids. *Journal of Molecular Biology* 2008, 375, 581-594.
17. Rokhsana, D.; Dooley, D.M.; Szilagyi, R.K.; Systematic development of computational models for the catalytic site in galactose oxidase: impact of outer-sphere residues on the geometric and electronic structures; *Journal of Biological Inorganic Chemistry* 2008. 13:371-383.
18. Pandey, A.S.; Harris, T.V.; Giles, L.J.; Peters, J.W.; Szilagyi, R.K.; Dithiomethylether as a Ligand in the Hydrogenase H-Cluster; *Journal of the American Chemical Society* 2008, 130:4533-4540.

Targeting and Photodynamic Killing of a Microbial Pathogen Using Protein Cage Architectures Functionalized with a Photosensitizer

Peter A. Suci,[†,‡] Zachary Varpness,[‡,§] Eric Gillitzer,[‡,||] Trevor Douglas,*[,‡,§] and Mark Young*[,‡,||]

Department of Microbiology, Center for BioInspired Nanomaterials, Department of Chemistry & Biochemistry, and Department of Plant Sciences, Montana State University, Bozeman, Montana 59717

Received July 17, 2007. In Final Form: August 20, 2007

The selectivity of antimicrobial photodynamic therapy (PDT) can be enhanced by coupling the photosensitizer (PS) to a targeting ligand. Nanoplatforms provide a medium for designing delivery vehicles that incorporate both functional attributes. We report here the photodynamic inactivation of a pathogenic bacterium, *Staphylococcus aureus*, using targeted nanoplatforms conjugated to a photosensitizer (PS). Both electrostatic and complementary biological interactions were used to mediate targeting. Genetic constructs of a protein cage architecture allowed site-specific chemical functionalization with the PS and facilitated dual functionalization with the PS and the targeting ligand. These results demonstrate that protein cage architectures can serve as versatile templates for engineering nanoplatforms for targeted antimicrobial PDT.

Introduction

The limitations of conventional antimicrobial therapy are being exposed by the manifestation of two forms of microbial resistance: acquired antibiotic resistance exhibited at the single-cell level and the intrinsic resistance of microbial biofilm communities.[1,2] A promising alternative, photodynamic therapy (PDT), relies on photosensitizers (PS) which when activated by light produce reactive oxygen species (ROS). ROS react with accessible cell components, subverting their function and finally causing cell death.[3,4] Bacterial strains that have acquired resistance to conventional antimicrobials are susceptible to antimicrobial PDT.[5−7] In addition, antimicrobial PDT has been successfully used to control biofilms[8−10] and forms the basis for emerging adjuvant and alternative treatments for biofilm infections of the oral cavity.[11]

Collateral damage to host tissue is a substantial concern for all forms of PDT, and the capability to more specifically target microbes with PS would significantly expand the range of possible clinical applications.[12] Nanoplatforms provide a relatively large surface area that can be used to engineer the presentation of both the PS and targeting ligand to combine cell selectivity with photodynamic killing.[13] Self-assembling protein cage architectures offer an exceptionally versatile template for design of multifunctional nanoplatforms that optimize functional group presentation, and thus are ideally suited as delivery vehicles for selective PDT.[14−23] The unique subunit structure of protein cages enables genetic insertion of chemically reactive amino acids or biologically reactive peptides into the monomeric protein subunits. These are subsequently presented symmetrically over the entire outer (or inner) surface of the assembled protein cage.[24,25] The ability to use a combination of genetic and chemical modifications to engineer protein cages expands the possibilities for design of multifunctional nanoplatforms that can be used for selective targeting of PS. In addition, their nonspecific adsorption to mammalian cells is minimal,[25,26] probably due to their intrinsic negative surface charge at physiological pH.

Staphylococcus aureus is a prominent player in infections that are difficult to treat due to acquired resistance to antibiotics[27,28]

* To whom correspondence should be addressed. E-mail: myoung@montana.edu (M.Y.); tdouglas@chemistry.montana.edu (T.D.).

† Department of Microbiology.

‡ Center for BioInspired Nanomaterials.

§ Department of Chemistry & Biochemistry.

|| Department of Plant Sciences.

(1) Costerton, J. W.; Stewart, P. S.; Greenberg, E. P. *Science* **1999**, *284*, 1318−1322.

(2) Walsh, C. *Nature* **2000**, *406*, 775−781.

(3) Brown, S. B.; Brown, E. A.; Walker, I. *Lancet Oncol.* **2004**, *5*, 497−508.

(4) Hamblin, M. R.; Hasan, T. *Photochem. Photobiol. Sci.* **2004**, *3*, 436−450.

(5) Szpakowska, M.; Lasocki, K.; Grzybowski, J.; Graczyk, A. *Pharmacol. Res.* **2001**, *44*, 243−246.

(6) Wainwright, M.; Phoenix, D. A.; Gaskell, M.; Marshall, B. *J. Antimicrob. Chemother.* **1999**, *44*, 823−825.

(7) Wilson, M.; Yianni, C. *J. Med. Microbiol.* **1995**, *42*, 62−66.

(8) Wainwright, M.; Crossley, K. B. *Int. Biodeterior. Biodegrad.* **2004**, *53*, 119−126.

(9) Wilson, M. *Photochem. Photobiol. Sci.* **2004**, *3*, 412−418.

(10) Zanin, I. C. J.; Lobo, M. M.; Rodrigues, L. K. A.; Pimenta, L. A. F.; Hofling, J. F.; Goncalves, R. B. *Eur. J. Oral Sci.* **2006**, *114*, 64−69.

(11) Meisel, P.; Kocher, T. *J. Photochem. Photobiol. B-Biol.* **2005**, *79*, 159−170.

(12) Demidova, T. N.; Hamblin, M. R. *Int. J. Immunopathol. Pharmacol.* **2004**, *17*, 245−254.

(13) Wang, S. Z.; Gao, R. M.; Zhou, F. M.; Selke, M. *J. Mater. Chem.* **2004**, *14*, 487−493.

(14) Douglas, T.; Young, M. *Science* **2006**, *312*, 873−875.

(15) Wang, Q.; Kaltgrad, E.; Lin, T. W.; Johnson, J. E.; Finn, M. G. *Chem. Biol.* **2002**, *9*, 805−811.

(16) Khayat, R.; Tang, L.; Larson, E. T.; Lawrence, C. M.; Young, M.; Johnson, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18944−18949.

(17) Kramer, R. M.; Li, C.; Carter, D. C.; Stone, M. O.; Naik, R. R. *J. Am. Chem. Soc.* **2004**, *126*, 13282−13286.

(18) Okuda, M.; Kobayashi, Y.; Suzuki, K.; Sonoda, K.; Kondoh, T.; Wagawa, A.; Kondo, A.; Yoshimura, H. *Nano Lett.* **2005**, *5*, 991−993.

(19) Yamashita, I. *Thin Solid Films* **2001**, *393*, 12−18.

(20) Gilles, C.; Bonville, P.; Rakoto, H.; Broto, J. M.; Wong, K. K. W.; Mann, S. *J. Magn. Magn. Mater.* **2002**, *241*, 430−440.

(21) McMillan, R. A.; Howard, J.; Zaluzec, N. J.; Kagawa, H. K.; Mogul, R.; Li, Y. F.; Paavola, C. D.; Trent, J. D. *J. Am. Chem. Soc.* **2005**, *127*, 2800−2801.

(22) Wiedenheft, B.; Mosolf, J.; Willits, D.; Yeager, M.; Dryden, K. A.; Young, M.; Douglas, T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10551−10556.

(23) Resnick, D. A.; Gilmore, K.; Idzerda, Y. U.; Klem, M. T.; Allen, M.; Douglas, T.; Arenholz, E.; Young, M. *J. Appl. Phys.* **2006**, *99*.

(24) Flenniken, M. L.; Willits, D. A.; Brumfield, S.; Young, M. J.; Douglas, T. *Nano Lett.* **2003**, *3*, 1573−1576.

(25) Flenniken, M. L.; Willits, D. A.; Harmsen, A. L.; Liepold, L. O.; Harmsen, A. G.; Young, M. J.; Douglas, T. *Chem. Biol.* **2006**, *13*, 161−170.

(26) Uchida, M.; Flenniken, M. L.; Allen, M.; Willits, D. A.; Crowley, B. E.; Brumfield, S.; Willis, A. F.; Jackiw, L.; Jutila, M.; Young, M. J.; Douglas, T. *J. Am. Chem. Soc.* **2006**, *128*, 16626−16633.

(27) Rice, L. B. *Biochem. Pharmacol.* **2006**, *71*, 991−995.

(28) Stefani, S.; Varaldo, P. E. *Clin. Microbiol. Infect.* **2003**, *9*, 1179−1186.

10.1021/la7021424 CCC: $37.00 © 2007 American Chemical Society
Published on Web 10/19/2007

and is also one of the primary biofilm-forming microbes involved in persistent biofilm infections.[1] We previously demonstrated that the viral protein cage (cowpea chlorotic mottle virus: CCMV) could be used to target *S. aureus* cells with MRI contrast agent.[29] CCMV is a 28 nm diameter icosahedral plant virus[30] that has served as a template for implementing several novel approaches for spatial control of multiligand presentation.[31,32] Here, we show that a genetically modified CCMV nanoplatform functionalized with PS can be targeted to *S. aureus* cells using both electrostatic and complementary biological interactions and that the cell-bound PS functionalized protein cages induce light-activated killing under standard light fluence conditions.

Experimental Section

Genetic Modification of CCMV/K42R and Production of the Constructs in Plants. The salt stable mutant of cowpea chlorotic mottle virus in which lysine 42 was replaced by an argenine (CCMV/K42R) has been described previously.[30] With use of this construct as starting material, polymerase chain reaction based site-directed mutagenesis was used to produce S102C/K42R and S130C/K42R, in which serines (102 and 130, respectively) were replaced by cysteines. Complementary oligonucleotide primers were 5′ GTTGCTTCCCAGTGTTTGTGGTACCGTGAAATCCTGTG3′(S102C/K42R, plus strand), 5′CACAGGATTTCACGGTACCACAAACACTGGGAAGCAAC 3′ (S102C/K42R, minus strand); 5′GCTGTGGCCGACAATTGCAAAGATGTTGTCGC 3′ (S130C/K42R, plus strand) and 5′ GCGACAACATCTTTGCAATTGTCGGCCACAGC 3′ (S130C/K42R, minus strand). Plasmids containing the mutated sequences were first screened by digestion with unique restriction enzymes (KpnI for S102C/K42R and MfeI for S130C/K42R). The mutations in the CCMV coat proteins of S102C/K42R and S130C/K42R were confirmed by DNA sequencing. Constructs were produced in cowpea plants and the virus was isolated as previously described.[33] Purity was verified using size exclusion chromatography (SEC) (Superose 6, Amersham Biosciences) and dynamic light scattering (DLS).[34]

Synthesis of Functionalized PS. 5-Iodoacetoamino-1,10-phenathroline (phen-IA) was synthesized by modification of a previously published protocol.[35] A solution of 1,3-dicyclohexylcarbodiimide (5.29 g) and iodoacetic acid (4.76 g) in 50 mL of dry ethyl acetate was stirred for 3 h at room temperature. The resulting solution was filtered to remove the urea, dried by rotary evaporation, and redissolved in 25 mL of acetonitrile. This solution was added to 25 mL of acetonitrile containing 5-amino-1,10-phenanthroline (1.0 g) and stirred overnight at room temperature. The product was collected by centrifugation, washed with cold 5% sodium bicarbonate and water, and dried under vacuum.

Bis(2,2′-bipyridine)dichlororuthenium ($Ru(bpy_2)Cl_2$) was synthesized according to literature procedures.[36] Ruthenium chloride trihydrate ($RuC1_3{\cdot}3H_20$) (7.8 g), bipyridine (9.36 g), and lithium chloride (8.4 g) were refluxed in dimethylformamide (50 mL) for 8 h. The reaction was cooled to room temperature, 250 mL of acetone was added, and the solution was stored at 4 °C overnight. The resultant product was filtered, washed with water and ether, and dried by suction.

$Ru(bpy_2)$phen-IA was synthesized by modification of a previously published protocol.[37] $Ru(bpy)_2Cl_2$ (0.7 g) and phen-IA (0.5 g) were refluxed in 50 mL of MeOH for 3 h with stirring. The solution was filtered and the product was precipitated by the addition of a concentrated aqueous solution of ammonium hexafluorophosphate. The orange solid was collected by filtration and washed with cold water and ether and dried in a desiccator.

PS Conjugation to S102C/K42R and S130C/K42R To Produce S102C/K42R-PS and S130C/K42R-PS. Genetically modified CCMV (S102C/K42R or S130C/K42R) (1.5 mg mL^{-1}, 75 μM monomer subunit) was dialyzed into 50 mM HEPES (pH 6.5) overnight and then transferred into fresh dialysis buffer (200 mL). Nitrogen was bubbled into the dialysis buffer for 1 h. 40 μL of a 5.5 mM solution of the PS ($Ru(bpy_2)$phen-IA) in DMS was added to 1.2 mL of the CCMV solution. This was stirred in the dark at room temperature for 5 h and then dialyzed into 50 mM HEPES (pH 6.5). SEC was used for further purification. Fractions eluting from within the CCMV peak, the position of which was predetermined using unlabeled CCMV, were combined. These were further purified by dialysis into 100 mM sodium acetate buffer (pH 4.8) and concentrated to 1.0 mg mL^{-1} protein using ultrafiltration (Amicon). Integrity of the labeled virus was confirmed by DLS and TEM as described previously.[34] UV−visible spectroscopy was used to determine the molar ratio of PS to monomer subunit. PS-conjugated S102C/K42R and S130C/K42R are referred to as S102C/ K42R-PS and S130C/K42R-PS, respectively.

Biotinylation of S130C/K42R-PS To Produce S130C/K42R-PS-B. S130C/K42R-PS was biotinylated by reaction with a 0.5 mM solution of sulfosuccinimidyl-6′-(biotinamido)-6-hexanamido hexanoate (Pierce) with 1.0 mg mL^{-1} of S130C/K42R-PS (50 μM monomer subunit) in 50 mM HEPES buffer, 150 mM sodium chloride at pH 7.1 at room temperature in the dark with stirring for 30 min. The reaction was terminated by exchange into 1 mM sodium acetate buffer at pH 4.8 using SEC. Fractions eluting from within the CCMV peak, the position of which was predetermined using unlabeled CCMV, were combined, further purified by dialysis into 100 mM sodium acetate buffer (pH 4.8), and concentrated to 1.8 mg mL^{-1} CCMV using ultrafiltration (Centricon Microcon YM-10). Integrity of the labeled virus was confirmed by DLS and TEM. Association of the biotin functional groups with the biotinylated product (S130C/K42R-PS-B) was initially tested by a dot blot assay using alkaline phosphatase conjugated antibiotin antibody (A-7064, Sigma-Alrich Co.).

Liquid Chromatography/Electrospray Mass Spectrometry. Liquid chromatography/electrospray mass spectrometry (LC/MS) was performed on a QToF Micro instrument (Waters). CCMV injected at 0.5−1.0 mg mL^{-1} (1−2 μL) was eluted from a Thermoelectric Biobasic size exclusion column (250 × 1 (mm)) in 80% isopropyl alcohol and 0.1% formic acid. The virus disassembled into monomers during the ionization process that were analyzed by the detector. Mass spectra of the multiply charged ions were deconvoluted using instrument software to produce a representation of monomer mass versus intensity.

Cell Culturing. ATCC strain 12598 (Cowan I) was maintained at −80 °C in 20% glycerol, 2% peptone. A solid medium (nutrient agar, Difco 0001) was streaked for single colonies. A single colony was used to inoculate 5 mL of a broth culture (nutrient broth, Difco 0003) and placed on a shaker at 37 °C at 280 rpm overnight. Fresh broth (5 mL) was inoculated from the starter culture to achieve an $OD_{600\ nm}$ of 0.250 and incubated as above for 3 h. This procedure resulted in a culture that had undergone about 3 doublings in the fresh medium and was in exponential phase (Supporting Information, Figure 1S). The Cowan I strain used in this work is a BSL2 pathogen. BSL2 practices were followed in accordance with the guidelines established at Montana State University.

Electrostatic Targeting of S102C/K42R-PS to Cells. Cells were cultured at 37 °C in batch in nutrient broth, harvested in exponential phase, and resuspended in phosphate buffered saline (PBS) (10 mM

(29) Suci, P. A.; Berglund, D. L.; Liepold, L.; Brumfield, S.; Pitts, B.; Davison, W.; Oltrogge, L.; Hoyt, K. O.; Codd, S.; Stewart, P. S.; Young, M.; Douglas, T. *Chem. Biol.* **2007**, *14*, 387−398.
(30) Speir, J. A.; Bothner, B.; Qu, C. X.; Willits, D. A.; Young, M. J.; Johnson, J. E. *J. Virol.* **2006**, *80*, 3582−3591.
(31) Gillitzer, E.; Suci, P.; Young, M.; Douglas, T. *Small* **2006**, *2*, 962−966.
(32) Klem, M. T.; Willits, D.; Young, M.; Douglas, T. *J. Am. Chem. Soc.* **2003**, *125*, 10806−10807.
(33) Bancroft, J. B.; Hiebert, E. *Virology* **1967**, *32*, 354.
(34) Basu, G.; Allen, M.; Willits, D.; Young, M.; Douglas, T. *J. Biol. Inorg. Chem.* **2003**, *8*, 721−725.
(35) Chen, C. H. B.; Sigman, D. S. *Science* **1987**, *237*, 1197−1201.
(36) Sullivan, B. P.; Salmon, D. J.; Meyer, T. J. *Inorg. Chem.* **1978**, *17*, 3334−3341.
(37) Castellano, F. N.; Dattelbaum, J. D.; Lakowicz, J. R. *Anal. Biochem.* **1998**, *255*, 165−170.

sodium phosphate, 100 mM sodium chloride, pH 7.0) to an OD_{600nm} of 0.2 (approximately 1.5×10^7 colony forming units (CFU) per mL). To form the poly-L-lysine (PLL) interlayer, the cell suspension was exposed to 6.25 μg mL^{-1} PLL (P1274, Sigma-Aldrich Co.) for 5 min. These cells were then exposed to S102C/K42R-PS (2 μM equivalent PS concentration).

Targeting of Cells with S130C/K42R-PS-B Using Complementary Biological Interactions. We modified previously published protocols to react biotinylated antibody with the cells and then react streptavidin (StAv) to the antibody-coated cells.[38,39] Exponential phase cells were pelleted at 4500Xg for 10 min and resuspended in 5 mL of PBS at an OD_{600nm} of 0.3. The cell suspension was incubated for 30 min on ice with a 1:100 dilution of biotinylated antiprotein A (SpA) monoclonal antibody (anti-SpA mAb-B) (P3150 Sigma-Aldrich Co.). This cell suspension was washed once in PBS to achieve a 1:100 dilution of the antibody and incubated with 20 μg mL^{-1} StAv (Sigma-Aldrich Co.) for 30 min on ice. After two washing steps in PBS (a 1:400 dilution of the StAv) the OD_{600nm} of the cell suspension was adjusted to 0.2. These cells were then exposed to S130C/K42R-PS-B (2 μM equivalent PS concentration).

Field Emission Scanning Electron Microscopy. PLL-coated Si ⟨100⟩ wafers (Virginia Semiconductor Inc., Fredericksburg, VA) were exposed to cells fixed in 3% gluteraldehyde in PBS for 1 h. The coupons were rinsed twice in Nanopure water and dried under a stream of nitrogen. The coupons with adsorbed cells were coated with a thin film of iridium by exposing the sample for 15 s at 20 mA in a Emitech sputter coater. Cells were viewed with a Supra 55VP FESEM (Zeiss) using the Inlens detector at 1 kV and 3 mm working distance.

Evaluation of Photodynamic Killing. For evaluation of photodynamic killing at the low fluence rate, cell suspensions were distributed in wells of a 96-microtiter well plate (200 μL) immediately after addition of S102C/K42R-PS and were exposed to a bank of LEDs (peak emission at 470 nm) in a temperature-controlled light box (20 °C) at a fluence rate of 0.763 mW cm^{-2} for 3.5 h, yielding a total fluence of 9.6 J cm^{-2}. For evaluation of photodynamic killing at the more standard fluence rate (46 mW cm^{-2}), cell suspensions were distributed into wells formed in a Teflon block (100 μL/well) 20 min after addition of S102C/K42R-PS or S130C/K42R-PS and illuminated by light from a mercury vapor lamp coupled to a Nikon Eclipse E600 microscope using a B2A filter block (excitation 450−490 nm; emission >515 nm). The stage height was adjusted so that the surface of the cell suspension was beyond the focal point of the 20× objective, 17 mm from the surface of the objective lens, to achieve a circular 7 mm spot that covered the surface of the cell suspension in the well. Fluence rate was determined using a photometer. Times of illumination were 5, 10, and 20 min to achieve total fluences of 13.8, 27.6, and 55.2 J cm^{-2}, which are in the standard range for antimicrobial PDT.[40] Viable cells numbers were evaluated by enumeration of CFU on a solid medium (nutrient agar) after incubation at 37 °C for 16 h.

Optical Microscopy. The red fluorescence from the PS associated with targeted cells was observed at 1000× (100× objective and 10× ocular) using the same microscope and filter set as used to illuminate cells for photodynamic killing at standard fluence rates. The fluorescence emission under these conditions is faint and rapidly photobleaches. However, the red fluorescence from targeted cells was clearly visible. Images were acquired with an Olympus Camedia camera coupled to the microscope.

Results and Discussion

As a delivery vehicle for PS, we modified a genetic construct (K42R) of CCMV that is exceptionally stable (referred to as CCMV/K42R). Site-specific mutagenesis was used to replace serines with reactive cysteines at surface-exposed sites on each of the 180 (20 kDa) protein monomer subunits that form the assembled CCMV/K42R protein cage. These served as attachment points for the PS (Figure 1). The PS, a ruthenium complex, was modified to react selectively with the sulfhydryls of the surface-exposed cysteines on these genetic constructs (S102C/K42R and S130C/K42R) by functionalizing the phenanthroline moiety with an iodoacetamide group. PS-conjugated constructs are referred to as S102C/K42R-PS and S130C/K42R-PS, respectively. UV−visible absorption spectroscopy was used to establish the association of the PS with each construct and to determine the molar ratio of bound PS to monomeric subunits (estimated as 0.520 for S102C/K42R-PS and 0.380 for S130C/K42R-PS) (see Supporting Information, Figure 2S). In addition, LC/MS was used to confirm covalent linkage of the PS to the CCMV constructs (see Supporting Information, Figure 3S).

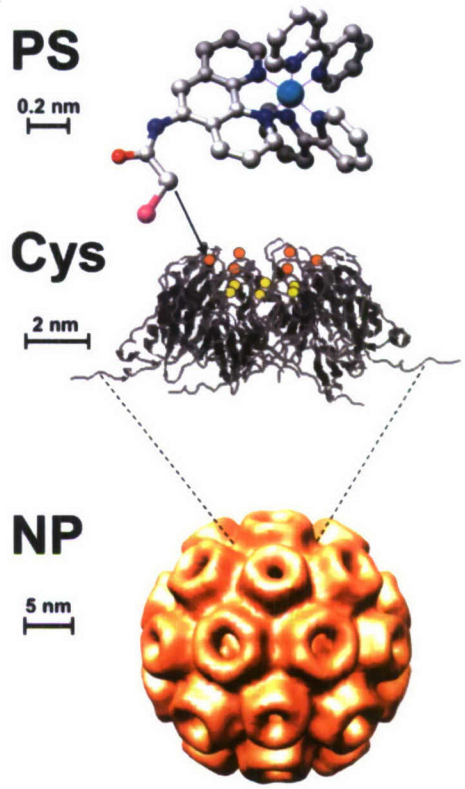


Figure 1. Genetically modified CCMV nanoplatform (NP) was functionalized with the PS (Ru(bpy$_2$)phen-IA) at surface-exposed sites. The positions of the sulfhydryl groups of the cysteines (Cys) added into the two genetic constructs are shown for a hexamer of protein subunits: S102C/K42R (orange) and S130C/K42R (yellow); the position of these sulfhydryls on the pentamers is similar. The PS was modified to promote reaction with exposed sulfhydryls by adding an iodoacetamide group to the phenanthroline moiety.

We have previously used both electrostatic and complementary biological interactions to direct the assembly of functionalized CCMV into organized films.[41] These same two strategies were used to target PS-conjugated nanoplatforms to the cells (Figure 2a). Cationic polymers, including poly-L-lysine (PLL), have been used to enhance the activity of PS.[42,43] In general, cell walls tend

(38) Wann, E. R.; Fehringer, A. P.; Ezepchuk, Y. V.; Schlievert, P. M.; Bina, P.; Reiser, R. F.; Hook, M. M.; Leung, D. Y. M. *Infect. Immun.* **1999**, *67*, 4737−4743.

(39) Yarwood, J. M.; McCormick, J. K.; Schlievert, P. M. *J. Bacteriol.* **2001**, *183*, 1113−1123.

(40) Jori, G.; Fabris, C.; Soncin, M.; Ferro, S.; Coppellotti, O.; Dei, D.; Fantetti, L.; Chiti, G.; Roneucci, G. *Lasers Surg. Med.* **2006**, *38*, 468−481.

(41) Suci, P. A.; Klem, M. T.; Arce, F. T.; Douglas, T.; Young, M. *Langmuir* **2006**, *22*, 8891−8896.

(42) Hamblin, M. R.; O'Donnell, D. A.; Murthy, N.; Rajagopalan, K.; Michaud, N.; Sherwood, M. E.; Hasan, T. *J. Antimicrob. Chemother.* **2002**, *49*, 941−951.

(43) Tegos, G. P.; Anbe, M.; Yang, C. M.; Demidova, T. N.; Satti, M.; Mroz, P.; Janjua, S.; Gad, F.; Hamblin, M. R. *Antimicrob. Agents Chemother.* **2006**, *50*, 1402−1410.

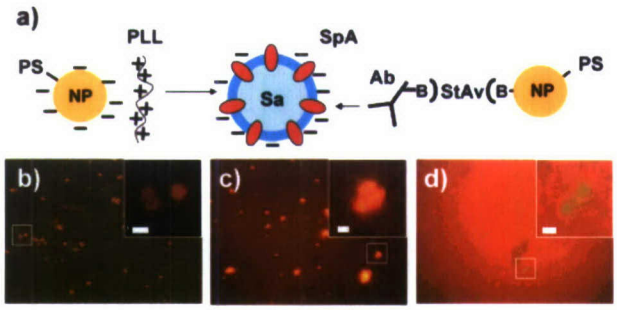


Figure 2. Targeting of PS-functionalized CCMV nanoplatforms to *S. aureus* cells. (a) Targeting strategies using electrostatic (left) or complementary biological interactions (right). Abbreviations are (from left to right): photosensitizer (PS); nanoplatform (NP); poly-L-lysine (PLL); *S. aureus* cell (Sa); protein A (SpA); anti-SpA-mAb-B (Ab), biotin (B) and streptavidin (StAv). (b, c, d) Epi-fluorescence images: (b) Cells targeted with S102C/K42R-PS via electrostatic interactions (exposure time, 0.5 s); (c) cells targeted with S130C/K42R-PS via complementary biological interactions (exposure time, 0.5 s); (d) cells mixed with S130C/K42R-PS (nontargeted) (exposure time, 1 s); the cells in (d) appear green due to faint intrinsic fluorescence; the red background originates from the PS-conjugated nanoplatform in solution. Scale bars in the magnified inserts are 1 μm.

to carry a net negative charge. Thus, we predicted that electrostatic interactions could provide a simple method to nonspecifically target the nanoplatform to *S. aureus*. In addition, PLL has intrinsic antimicrobial properties,[44] suggesting that an interlayer of PLL on the cell wall would enhance the activity of a PS-functionalized nanoplatform. Whereas electrostatic interactions can be used to mediate relatively nonspecific targeting to cells, complementary biological interactions offer a means to selectively target PS to specific microbes.[45−47] S130C/K42R-PS was specifically targeted to protein A (SpA) presented in the cell wall of *S. aureus* by using streptavidin (StAv) to couple biotinylated anti-SpA monoclonal antibody (anti-SpA mAb-B) to biotinylated nanoplatform (Figure 2a). SpA binds to the Fc region of IgG and may play a role in immune avoidance.[48] We used flow cytometry to characterize expression of SpA in the *S. aureus* ATCC 12598 strain cultured under our conditions (see Supporting Information, Figure 4S).

A biotin targeting ligand was added to the endogenous lysines of S130S/K42R-PS to produce S130S/K42R-PS-B. The S130C/K42R construct was chosen for dual functionalization with PS and the biotin functional group based on the consideration of minimizing possible interference with binding of StAv to the biotin targeting ligand due to steric hindrance from the relatively large PS functional group. Preliminary results indicated that one of the most reactive endogenous lysines (K87) is located near the pseudo 3-fold axis of the CCMV viral capsid which is well-spaced from the reactive cysteines on S130C/K42R. Covalent linkage of the biotin functional groups into the S130C/K42R-PS-B construct was confirmed by LC/MS (see Supporting Information, Figure 3S).

The presence of PS-conjugated nanoplatform on targeted cells was first verified by epi-fluorescence microscopy (Figures 2b−2d). Targeted cells appeared red while cells exposed to PS-conjugated nanoplatform without targeting produced a faint green intrinsic fluorescence which was only visible in images acquired at sufficiently long exposure times. The red fluorescence from cells targeted using electrostatic interactions was clearly more faint than the fluorescence from cells targeted using complementary biological interactions (antibody-SpA). The red background in the images originates from free PS-conjugated nanoplatform in solution. The presence of PLL probably precipitates a portion of the nanoplatform in solution, causing the diffuse background in Figure 2b to be fainter than that in Figure 2c.

The arrangement of targeted nanoplatform on the cell wall was characterized directly by FESEM (Figure 3). The FESEM images indicated that the cell wall associated nanoplatforms maintained their integrity. Targeting via electrostatic interactions produced more sparse coverage than targeting via complementary biological interactions and, in addition, targeting via complementary biological interactions produced a more uniform coverage. Twenty images were acquired of cells targeted via this latter scheme. The image presented in Figure 3c is representative of the high density of coverage exhibited by cells targeted with complementary biological interactions (see Supporting Information, Figure 5S). The level of nonspecific binding of nontargeted PS-functionalized nanoplatform (S102C/K42R-PS) shown in Figure 1a is representative of the low level observed for S130C/K42R-PS as well (see Supporting Information, Figure 6S).

Cell populations targeted with the PS-conjugated CCMV constructs via both electrostatic and complementary biological interactions were photodynamically inactivated to a significant level (Figures 4−6). Cells targeted with S102C/K42R-PS via electrostatic interactions mediated by PLL were photodynamically inactivated at a very low fluence rate (Figure 4) as well as at a more standard fluence rate and standard total fluence levels (Figure 5). Without the PLL interlayer, PS conjugated to the nanoplatform was almost ineffectual in killing cells: nontargeted PS-conjugated nanoplatoforms (S102C/K42R-PS and S130C/K42R-PS) did not induce appreciable photodynamic killing under conditions in which free PS in solution reduced viable cell numbers by 5 orders of magnitude (see Supporting Information, Figure 7S). Results presented in Figures 4 and 5 show that a PLL interlayer on the cell wall rendered the level of killing activity of PS conjugated to the nanoplatform to within the range of free PS. These results demonstrate that PS at concentrations that are normally cidal upon light activation is innocuous when the PS is conjugated to the CCMV nanoplatform (but not targeted to cells), while targeted PS-conjugated CCMV can be used to obtain high local PS concentrations which are cidal.

The killing activity of S102C/K42R-PS targeted via electrostatic interactions is likely attributable to a combination of membrane disruption by PLL and ROS production by PS located proximal to the cell membrane. Under dark conditions, PLL at the concentration used for these experiments (6.25 μg mL^{-1}) reduced the viable cells by about half an order of magnitude, probably by compromising the cell membrane (Figures 4 and 5c). Association of S102C/K42R-PS with cells pre-exposed to PLL significantly reduced the number of viable cells compared to exposure to PLL alone for experiments at the low fluence rate (8 independent replicates, paired t-test, p value 0.006) (Figure 4). This same trend was observed at a more standard fluence rate and standard total fluence levels (Figures 5a and 5c). The association of PLL with the cell wall also enhanced the killing activity of free PS in solution (compare Figure 5b and Figure 6c) as anticipated from previous results.[42,43] The dependence of

(44) Delihas, N.; Riley, L. W.; Loo, W.; Berkowitz, J.; Poltoratskaia, N. *FEMS Microbiol. Lett.* **1995**, *132*, 233−237.
(45) Embleton, M. L.; Nair, S. P.; Heywood, W.; Menon, D. C.; Cookson, B. D.; Wilson, M. *Antimicrob. Agents Chemother.* **2005**, *49*, 3690−3696.
(46) Embleton, M. L.; Nair, S. P.; Cookson, B. D.; Wilson, M. *Microb. Drug Resist. (Larchmont, NY, U.S.)* **2004**, *10*, 92−97.
(47) Embleton, M. L.; Nair, S. P.; Cookson, B. D.; Wilson, M. *J. Antimicrob. Chemother.* **2002**, *50*, 857−864.
(48) Patel, A. H.; Nowlan, P.; Weavers, E. D.; Foster, T. *Infect. Immun.* **1987**, *55*, 3103−3110.

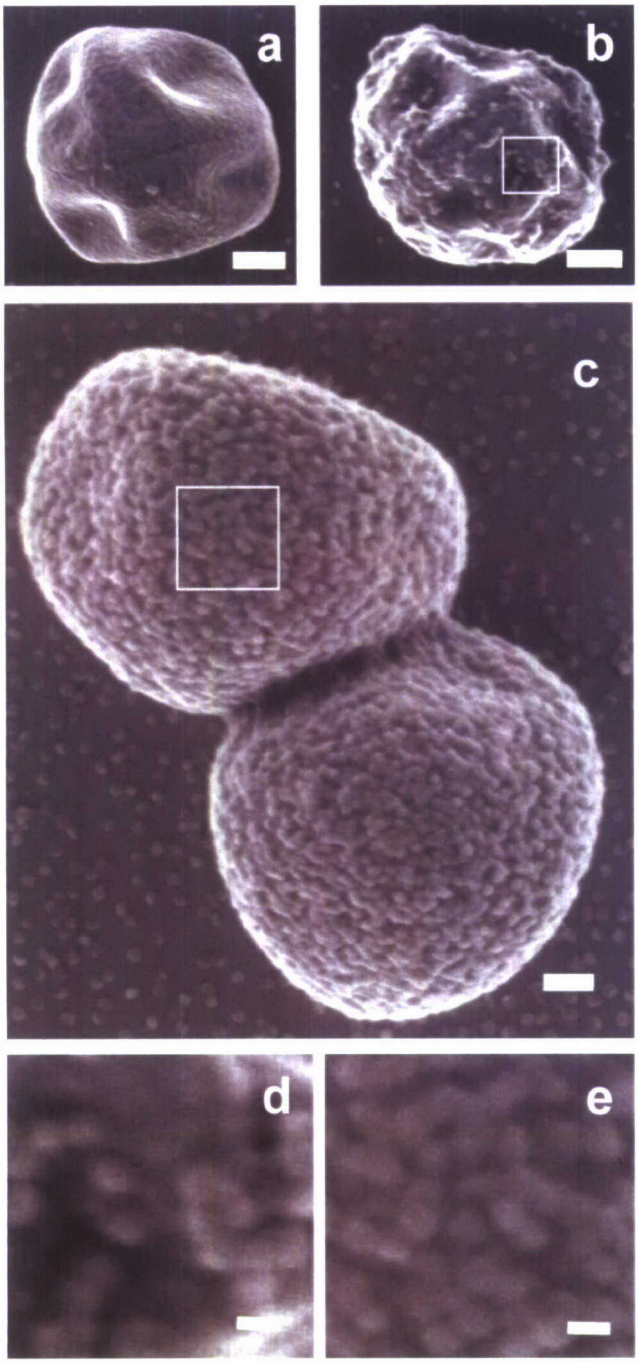


Figure 3. FESEM images showing the arrangement of targeted nanoplatforms on the cell wall: (a) cell exposed to S102C/K42R with no targeting; (b) cell targeted with S102C/K42R-PS via electrostatic interactions; (c) cells targeted with S130C/K42R-PS via complementary biological interactions; (d) magnified view of the region indicated in (b); (e) magnified view of the region indicated in (c). Scale bars: (a ,b) 200 nm; (c) 100 nm; (d, e) 28 nm (the diameter of CCMV). Note: CCMV in solution adsorbs strongly to the positively charged substratum used for FESEM visualization.[49]

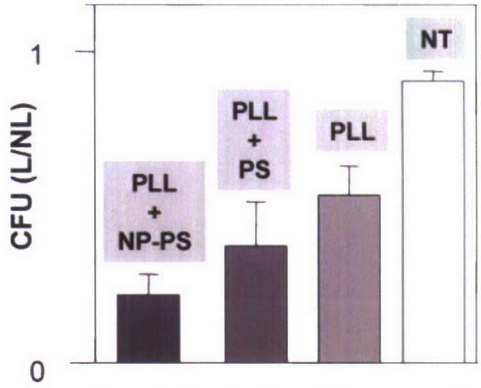


Figure 4. Photodynamic inactivation of *S. aureus* cells targeted with PS-functionalized nanoplatform (NP-PS) via electrostatic interactions at a fluence rate of 0.763 mW cm^{-2} (total fluence of 9.6 J cm^{-2}). Data are presented as the ratio of CFU for illuminated cells to CFU for cells maintained in the dark. Error bars are standard errors for the mean of eight independent replicates: PLL + NP-PS: cells exposed to PLL and targeted with S102C/K42R-PS at 2 μM equivalent PS concentration; PLL + PS: cells exposed to PLL and the photosensitizer (2 μM); PLL: cells exposed only to PLL; NT: no treatment. A paired t-test indicated that the difference between the means for the pairs (PLL) and (PLL + NP-PS) were significant at the 1% level, whereas the difference between the means for the pairs (PLL + PS) and (PLL + NP-PS) were not significant at the 10% level. (The differences between the mean for the untreated condition and the means for the other three conditions were each significant at the 1% level.)

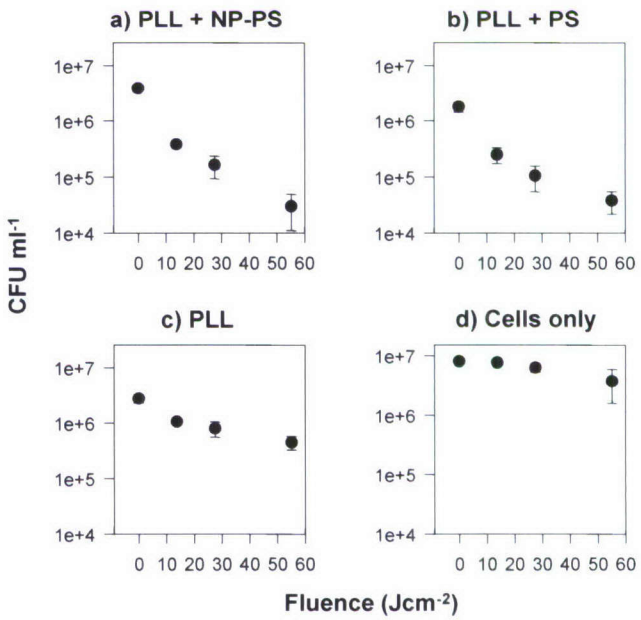


Figure 5. Photodynamic inactivation of *S. aureus* cells targeted with PS-functionalized nanoplatform (NP-PS) via electrostatic interactions at a fluence rate of 46 mW cm^{-2}. Data are presented as CFU mL^{-1} recovered after treatment. Error bars are standard errors for the mean of two independent replicates. (a) Cells exposed to PLL and targeted with S102C/K42R-PS at 2 μM equivalent PS concentration; (b) cells exposed to PLL and the photosensitizer (2 μM); (c) cells exposed only to PLL; (d) no treatment.

the killing activity on light fluence in the absence of PS exhibited in Figure 5c, and much less prominently in Figure 5d, suggests that there were intrinsic PS associated with the cell wall whose killing activity was also enhanced by PLL.[50,51] The results suggest that the killing activities conferred by a cationic polymer or peptide and a PS both conjugated to a nanoplatform would combine to produce significant levels of cell inactivation.

S130C/K42R-PS-B targeted to *S. aureus* cells using complementary biological interactions produced levels of photodynamic

(49) Suci, P. A.; Klem, M. T.; Douglas, T.; Young, M. *Langmuir* **2005**, *21*, 8686−8693.

(50) Nitzan, Y.; Salmon-Divon, M.; Shporen, E.; Malik, Z. *Photochem. Photobiol. Sci.* **2004**, *3*, 430−435.

(51) Kafala, B.; Sasarman, A. *Gene* **1997**, *199*, 231−239.

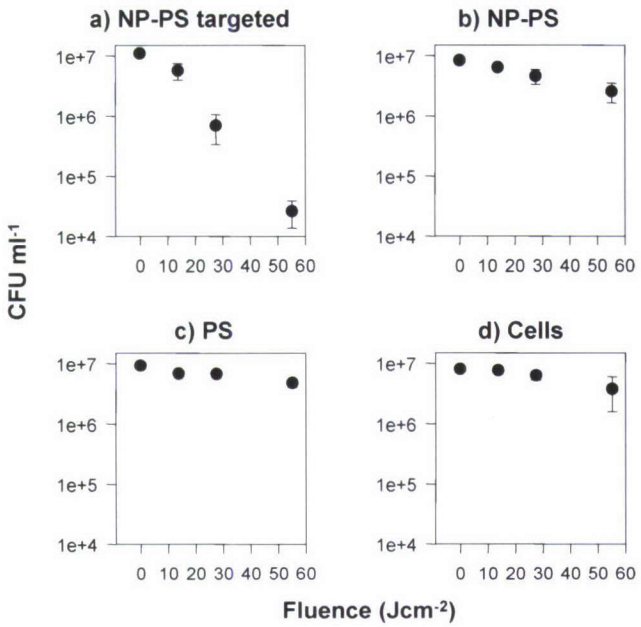


Figure 6. Photodynamic inactivation of *S. aureus* cells targeted with PS-functionalized nanoplatform (NP-PS) via complementary biological interactions at a fluence rate of 46 mW cm^{-2}. Data are presented as CFU mL^{-1} recovered after treatment. Error bars are standard errors for the mean of three independent replicates. (a) Cells targeted with S130C/K42R-PS-B at 2 μM equivalent PS concentration; (b) cells exposed to S130C/K42R-PS at 2 μM equivalent PS concentration (nontargeted); (c) cells exposed to 2 μM PS; (d) cells targeted with anti-SpA-mAb-B and StAv but not with the PS-functionalized nanoplatform. An unpaired t-test indicated that the means for the targeted cells were significantly different from means for all other conditions at fluences of 27.6 and 55.2 J cm^{-2} at the 5% level.

killing that were significantly greater than the free PS in solution (2 μM) and nontargeted S130C/K42R-PS at an equivalent PS concentration (Figure 6). Data presented in Figure 6 are the means and standard error for three independent experiments. The reduction in viable cells at 27.6 and 55.2 J cm^{-2} is significantly greater at the 5% level of confidence for cells targeted with S130C/K42R-PS-B compared with cells exposed to S130C/K42R-PS (not targeted) and cells exposed to PS alone (unpaired t-test).

The level of killing produced in cells targeted via complementary biological interactions (Figure 6a) was approximately the same as that achieved by targeting SpA in *S. aureus* with an antibody conjugated to a PS under comparable conditions of both light fluence levels and equivalent PS concentration.[47] There were approximately 70 PS delivered per binding event for CCMV targeted via complementary interactions, while only about 2 PS per binding event were delivered using direct conjugation to the antibody. Therefore, the level of killing obtained with the targeted nanoplatform might be expected to exceed that obtained with PS-conjugated antibody if killing were limited by PS per binding event. One possibility is that the relatively large size of CCMV significantly reduced the proximity of PS to critical cell components. Proximity is essential for the activity of singlet oxygen which has a very short diffusion length (<50 nm) and is thought to be the primary ROS causing oxidative damage.[52] We have not attempted to construct a detailed model of binding of CCMV to SpA using our scheme (Figure 2), but approximate modeling indicates that the distance between PS coupled to CCMV that is bound to cells and the cell membrane was less than 50 nm (Supporting Information, Figure 8S).

The killing activity might have been limited by the distribution of SpA expression in the cell population, rather than the quantity of PS delivered per binding event. This was proposed as a primary constraint on killing in the case of targeting with PS-conjugated antibody.[47] The Cowan I strain used in our experiments is known to express SpA at a high level.[38] However, according to the flow cytometry analysis the level of expression of SpA in a small portion of the cell population was very low or negligible (see Supporting Information, Figure 4S). Our results indicate that the majority of these nontargeted cells would survive exposure to the fluence levels used in our experiments. The PS loading advantage offered by a targeted CCMV nanoplatform (compared to a smaller biomolecule such as an antibody) might have been more evident in a case in which epitope presentation in the cell wall was more sparse, but uniformly distributed among the entire cell population.

In addition to offering the possibility of delivering substantial PS per binding event, the relatively large surface area of nanoplatforms provides space for decoration with multiple targeting ligands. The ability to engineer placement of multiple targeting ligands on a PS delivery vehicle opens up the possibility of finely tuning selectivity. This capability could be used as a means to eliminate pathogens while preserving beneficial commensal microbes for treatment of oral infections,[11] or to ensure selective eradication of genetic or physiological variants of a pathogen causing a recalcitrant wound infection.[12] In the case of *S. aureus* there are at least 22 different cell wall anchored proteins that are involved in conferring virulence and could potentially be targeted with appropriately functionalized nanoplatforms.[53,54]

For topical applications the indirect targeting scheme used in our experiments would be feasible. Similar strategies, involving pretargeting with StAv, have been proposed even for intravenous applications.[55] Protocols for biotinylation of antibodies that preserve binding affinity of the active site are well-established, and our results demonstrate that site-specific dual labeling of CCMV genetic constructs with both a functional biotin and an active PS is manageable. In this context, nanoplatforms such as protein cage architectures provide an accessible template for implementing a modular approach to selective PDT. The advantage of this alternative approach is that strategies for targeting multiple ligands and optimizing PS loading per binding event could be developed independently and then coupled via the nanoplatform.

Conclusion

Our results demonstrate the feasibility of using PS-functionalized nanoplatforms based on genetically modified CCMV as targeted delivery vehicles for selective antimicrobial PDT. If the requirement for selectivity is not too stringent, electrostatic interactions can be used to target PS-functionalized CCMV to microbial pathogens. In this case relatively dilute concentrations of the cationic polymer significantly enhance the photodynamic killing. To achieve true selectivity, complementary biological interactions can be used. The simplest strategy in this case is to dual functionalize the nanoplatform with a biotin and the PS.

(52) Ochsner, M. *J. Photochem. Photobiol. B-Biol.* **1997**, *39*, 1−18.

(53) Marraffini, L. A.; DeDent, A. C.; Schneewind, O. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 192−221.

(54) Harris, L. G.; Foster, S. J.; Richards, R. G. *Eur. Cells Mater.* **2002**, *4*, 39−60.

(55) Wu, A. M.; Senter, P. D. *Nat. Biotechnol.* **2005**, *23*, 1137−1146.

The advantage of being able to use a combination of both chemical and genetic methods to control functional group presentation lends protein cage architectures a versatility not offered by most other nanoplatforms. This versatility can be exploited to pursue a modular approach to design vehicles for selective antimicrobial PDT in which PS loading and targeting functions can be independently optimized.

Acknowledgment. This work was funded by grants from NIH (R01 EB00432) and the Office of Naval Research for support of the Center for BioInspired Nanomaterials (19-00-R0006).

Supporting Information Available: Figures 1S−7S. This material is available free of charge via the Internet at http://pubs.acs.org.

LA7021424

Viral Capsids as MRI Contrast Agents

Lars Liepold,[1,2] Stasia Anderson,[3] Deborah Willits,[1,4] Luke Oltrogge,[1,2] Joseph A. Frank,[5] Trevor Douglas,[1,2*] and Mark Young[1,4]

Viral capsids have the potential for combined cell/tissue targeting, drug delivery, and imaging. Described here is the development of a viral capsid as an efficient and potentially relevant MRI contrast agent. Two approaches are outlined to fuse high affinity Gd^{3+} chelating moieties to the surface of the cowpea chlorotic mottle virus (CCMV) capsid. In the first approach, a metal binding peptide has been genetically engineered into the subunit of CCMV. In a second approach gadolinium-tetraazacyclododecane tetraacetic acid (GdDOTA) was attached to CCMV by reactions with endogenous lysine residues on the surface of the viral capsid. T_1 and T_2 ionic relaxivity rates for the genetic fusion particle were R1 = 210 and R2 = 402 $mM^{-1}s^{-1}$ (R2 at 56 MHz) and for CCMV functionalized with GdDOTA were R1 = 46 and R2 = 142 $mM^{-1}s^{-1}$ at 61 MHz. The relaxivities per intact capsid for the genetic fusion were R1 = 36,120 and R2 = 69,144 $mM^{-1}s^{-1}$ (R2 at 56 MHz) and for the GdDOTA CCMV construct were R1 = 2,806 and R2 = 8,662 $mM^{-1}s^{-1}$ at 61 MHz. The combination of high relaxivity, stable Gd^{3+} binding, and large Gd^{3+} payloads indicates the potential of viral capsids as high-performance contrast agents. Magn Reson Med 58:871–879, 2007. © 2007 Wiley-Liss, Inc.

Key words: MRI; contrast agent; viral capsids; cowpea chlorotic mottle virus; gadolinium

MRI is one of the most utilized imaging techniques in medicine since it is noninvasive and provides comparatively high-resolution information. The usefulness of the technique is augmented by the use of contrast agents that increase the rate of water proton relaxation and can therefore increase contrast between tissues. Gadolinium (Gd^{3+}) chelates are commonly used as contrast agents in clinical settings (1,2). In general, there are two ways to improve the imaging sensitivity using contrast agents: either by increasing the relaxivity of water protons through direct interaction with the contrast agent or by targeted delivery of the agent to specific locations within the body.

Viral capsids are multimeric protein assemblies that form cage architectures and can be generally categorized as protein cages. Other, nonviral protein cages include heat shock proteins, ferritins, and vault ribonucleoprotein particles, among others. These protein cages can serve as robust synthetic platforms that are chemically and genetically malleable and can be readily modified. Previous studies have explored the use of protein cages as therapeutic or imaging delivery agents (3–5). Cell targeting has been achieved by utilizing capsids with natural affinities for cellular receptors or by chemically linking peptides or antibodies to protein cage architectures (6,7). In addition, targeted protein cages incorporating a therapeutic payload (doxorubicin) have been constructed, demonstrating the multifunctional capacity for biomedical applications (8).

Protein cages and more specifically viral capsids, have the potential to serve as extremely efficient contrast agents for the following reasons: 1) viral capsids are large, commonly between 18–100 nm in diameter, and relatively rigid molecular structures with large rotational correlation times, resulting in increased relaxivity rates; 2) viral capsids can serve as robust platforms onto which multiple functional motifs can be added through genetic or chemical modifications (9–16). These modifications could potentially result in the attachment of both Gd^{3+} binding and site-specific targeting functionalities; and 3) viral capsids can potentially carry hundreds (if not thousands) of Gd^{3+} ions and the contrast from an individual capsid will increase significantly with the number of Gd^{3+} ions it carries. Due to these three reasons, viral capsids have been investigated as MRI contrast agents, including the cowpea chlorotic mottle virus (CCMV) capsid with bound Gd^{3+} at endogenous metal bind sites and the MS2 virus capsid with gadolinium diethylenetriamine pentaacetic acid (Gd-DTPA) chemically attached (4,17). Dendrimers, liposomes, and other supramolecular structures maintain properties 2 and 3 mentioned above and therefore have also been developed as potential contrast agents (18,19).

In this study, potential MRI contrast agents based on the multivalent protein cage architecture of the CCMV capsid have been developed. The CCMV viral capsid is assembled from 180 identical protein subunits. The capsid forms an icosahedral architecture through noncovalent interactions between the subunits. A consequence of the cage-like architecture assembly is that amino acid residues are organized and presented in a repeating and symmetrical fashion over the capsid surface. Therefore, genetic and chemical modifications of amino acids within the subunit results in the presentation of the modification at all 180 symmetrically-related sites on the viral capsid. The exterior diameter of the CCMV capsid is 28 nm and the interior diameter is 22 nm (Fig. 1) (20).

CCMV has been shown to contain an endogenous metal binding site. In a previous study characterizing the metal binding by fluorescence resonance energy transfer (FRET), tryptophan residues proximal to the endogenous metal binding site were used to determine Tb^{3+}, Gd^{3+}, and Ca^{2+}

[1]Center for BioInspired Nanomaterials, Montana State University, Bozeman, Montana, USA.

[2]Department of Chemistry and Biochemistry, Montana State University, Bozeman, Montana, USA.

[3]Mouse Imaging Core, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA.

[4]Department of Plant Sciences, Montana State University, Bozeman, Montana, USA.

[5]Experimental Neuroimaging Section (Clinical Center), National Institutes of Health, Bethesda, Maryland, USA.

Grant sponsor: National Institutes of Health; Grant numbers: R21EB005364, R01 EB00432; Grant sponsor: Office of Naval Research; Grant number: N00014-04-1-0672.

*Correspondence to: Trevor Douglas, Center for BioInspired Nanomaterials, Montana State University, Bozeman, MT, 59717.
E-mail: tdouglas@chemistry.montana.edu

Received 4 December 2006; revised 19 April 2007; accepted 23 April 2007.
DOI 10.1002/mrm.21307
Published online in Wiley InterScience (www.interscience.wiley.com).

© 2007 Wiley-Liss, Inc.

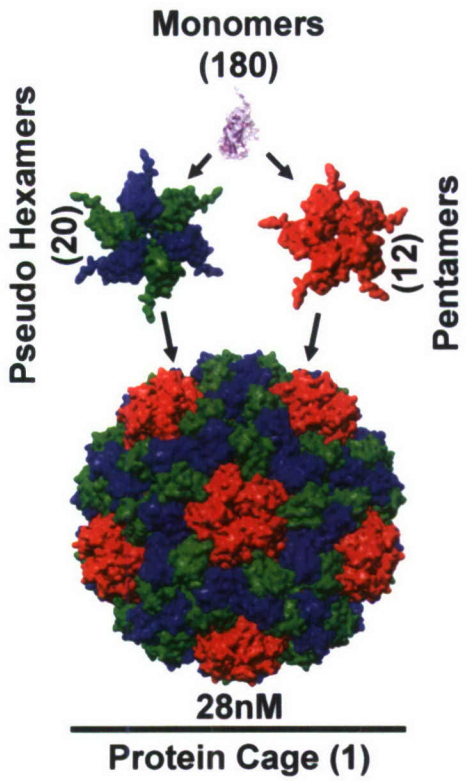


FIG. 1. The 28-nm viral capsid of CCMV is made of 180 individual subunit proteins (20 hexamers and 12 pentamers).

dissociation constants (K_d = 19 μM, 31 μM, and 2 mM, respectively) (21). Further studies indicated that Gd^{3+}-bound CCMV (CCMV-Gd) capsids exhibited high ionic relaxivity rates (R1 = 202, R2 = 376 $mM^{-1}s^{-1}$, at 61 MHz) and extremely high relaxivity rates per capsid (R1 = 28,482 and R2 = 53,016 $mM^{-1}s^{-1}$, 61 MHz) (17). However, Gd^{3+} binding to the endogenous sites was too weak for the CCMV-Gd to be useful as a clinical MRI contrast agent, since free Gd^{3+} is toxic in vivo (22).

The purpose of this study was to take advantage of the CCMV architecture while enhancing the Gd^{3+} binding constant of the lanthanide metal ion in the development of two nanoparticle contrast agents. The first approach was to genetically incorporate a nine-residue peptide sequence, from the Ca^{2+} binding protein calmodulin, as a genetic fusion to the N-terminus of the CCMV subunit (CCMV-CAL). Characterization of the metal binding to the genetically engineered capsid was undertaken using FRET analysis. The second approach was to covalently attach the clinically relevant contrast agent gadolinium-tetraazacyclododecane tetraacetic acid (GdDOTA) to reactive lysine residues on CCMV via an *N*-hydroxysuccinimide (NHS) ester coupling reaction (CCMV-DOTA). Both R1 and R2 relaxivity data were measured as a function of field strength for the genetic and chemically-modified viral capsid contrast agents.

MATERIALS AND METHODS

Engineering CCMV to Express the Metal Binding Sequence of Calmodulin

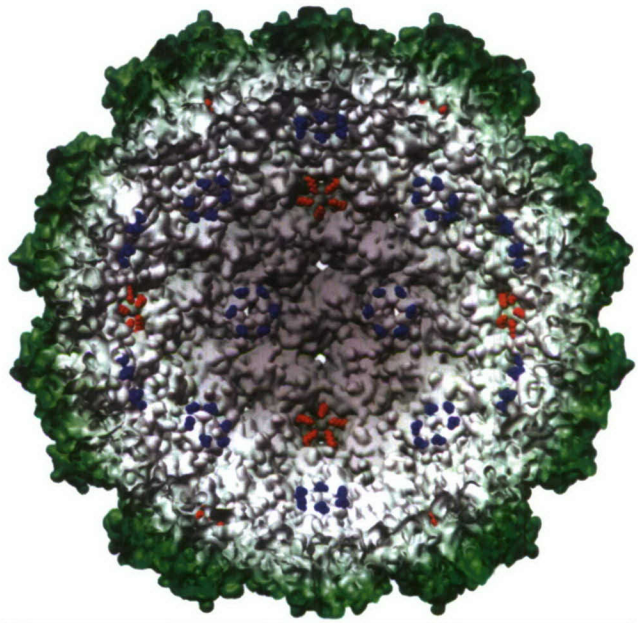
The SubE/R26C/K42R gene, a mutant of the coat protein of the CCMV capsid, cloned into the *Pichia pastoris* vector; pPICZA (Invitrogen) was used as the template (23). QuikChange site-directed mutagenesis (Stratagene) using the primer; (5′cgaggaattcatgtctacagacaaagatggtgatggatggttagaattcgaagagggtgggggcgaagagaacgaggagaacac3′), and its reverse complement was used to insert the calmodulin deoxyribonucleic acid (DNA) sequence into the N-terminus of the coding region of the capsid protein. The mutagenized vector from CCMV was confirmed by DNA sequencing (Applied Biosystems). This modified CCMV protein was called CCMV-CAL.

Expression and Purification of CCMV-CAL

The mutagenized CCMV capsid protein gene was expressed and purified in a *Pichia pastoris* heterologous protein expression system as previously described (23). High levels of coat protein expression were induced and yielded assembled viral protein capsids devoid of nucleic acid.

These viral capsids were purified to near homogeneity by lysis of cells, followed by ion exchange chromatography. Size-exclusion chromatography (SEC) was used to further purify the capsid and to eliminate any aggregates or subunit disassembly products potentially present in the samples (Superose 6; Amersham Biosciences; 50 mM 4-2-hydroxyethyl-1-piperazineethanesulfonic acid (HEPES), pH 6.5). Both ion exchange and SEC were performed on an Amersham Akta purifier FPLC.

The 12-residue binding sequence and three glycine residues, shown in (Fig. 2), replaced the original residues



	1	11
Unmodified:	MSTVGTGELT	EAQEEAAAEE
CCMV-CAL:	MSTDKDGDGW	LEFEEGGGEE

FIG. 2. Inside view of CCMV's viral capsid. Blue highlights are residue 27 in the six-fold environment. Red highlights are residue 42 in the five-fold environments. The first 20 amino acids are shown for both the unmodified viral subunit and genetically modified subunit. The underlined residues are responsible for metal binding.

4–18 of the SubE/R26C/K42R CCMV mutant. This replacement was confirmed at the protein level by liquid chromatography/mass spectrometry (LC/MS) of the purified protein (Agilent Technologies 1100 LC system coupled to an Esquire 3000 ion trap mass spectrometer, Bruker Daltonics). The theoretical mass was calculated by considering loss of the N-terminal methionine and acetylation of the second residue (serine). Protein concentration was determined by the absorbance at 280 nm (ϵ = 29280 $M^{-1}cm^{-1}$ for CCMV-CAL, ϵ values were calculated by inserting amino acid sequences into the ProtParam tool at http://ca.expasy.org/tools/protparam.html).

Synthesis of CCMV- DOTA Capsids

The following buffers were used in the synthesis of CCMV-DOTA; labeling buffer (100 mM HEPES, 100 mM NaCl, pH 7.2), and storage buffer (100 mM HEPES, 100 mM NaCl, pH 6.5). A lysine reactive form (NHS-ester) of the metal chelator DOTA was used in the synthesis (Macrocyclics, B-280). Figure 3 outlines the general reaction scheme. The K42R mutant of the CCMV (0.5–3 mg/ml, 25–150 μM subunit) virus particle was purified from infected plants (as previously described) and dialyzed into 500 mL of labeling buffer for 3 h (24). The concentration of plant virus was calculated by multiplying the absorbance at 260 nm by 6.4 to yield a concentration of CCMV subunit in μM units. A concentration of 1 to 2 mg/ml (~50 to 100 μM subunit) was typically used in the reaction. A 20× (mole:mole) of NHS-ester DOTA was added next. The pH was maintained at 7.0 by additions (1–5 μl) of 0.5 M NaOH. The reaction mixture was monitor by LC/MS (both Standard and Nano Aquity LC systems and both Q-Tof Micro and Q-Tof Premier mass spectrometers were used). SEC and reverse-phase separation techniques were used. The deconvolution program MaxEnt1 (Waters) was used to determine the percent of subunits with DOTA covalently linked to them. Equation [1] was used to approximate the average labeling of subunits with DOTA, where D is the number of DOTAs attached to the subunit and I_D is the intensity of the ion corresponding to a subunit with D DOTAs attached.

$$\frac{AverageDOTA}{Subunit} = \frac{\sum_{0}^{D} DI_D}{\sum_{0}^{D} I_D}. \qquad [1.1]$$

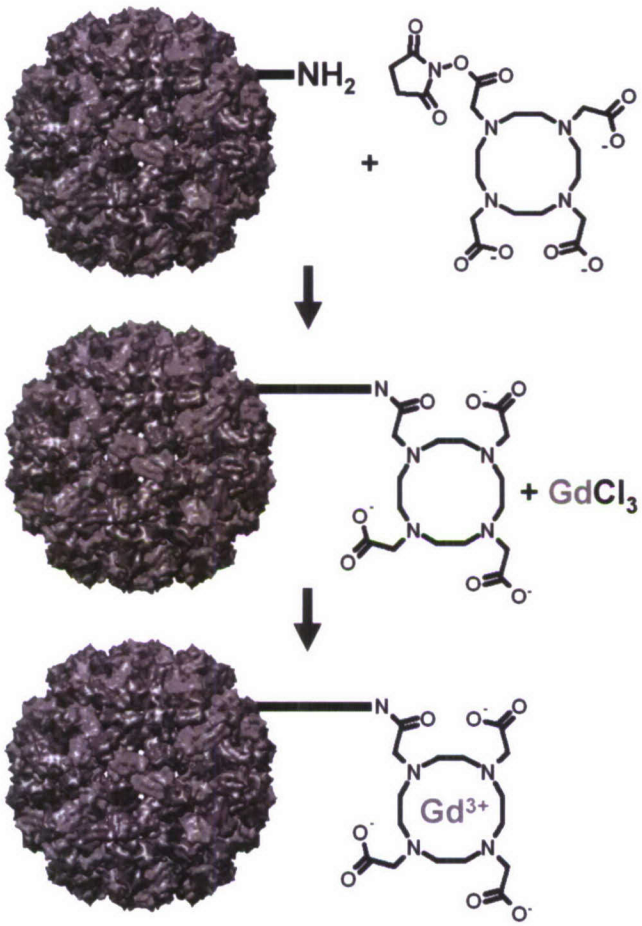


FIG. 3. The reaction scheme to attach DOTA-Gd^{3+} to the CCMV viral capsid. Endogenous lysines on the viral capsid are reacted to a DOTA/NHS conjugation. Next, $GdCl_3$ is added to produce a viral capsid conjugated with Gd^{3+} ions.

The reaction was allowed to proceed for 2 h at 25°C. At that point, LC/MS analysis revealed that the majority of NHS-DOTA reactant was hydrolyzed. The reaction was repeated (with additions of NHS-ester DOTA) until there was, on average, one DOTA covalently attached per subunit. Unreacted DOTA was removed by dialyzing the reaction mixture into labeling buffer. Next the CCMV-DOTA was dialyzed into labeling buffer with 10× $GdCl_3$ (moles of Gd^{3+}:moles of subunit) and a pH of 7.0. The progression of the metal loading onto the CCMV-DOTA was monitored by LC/MS. Free Gd^{3+} was separated from the CCMV-DOTA-Gd by dialyzing the labeled capsid into storage buffer that contained 5 mM ethylene diamine tetraacetic acid (EDTA). Then multiple dialysis steps were performed into storage buffer. Alternatively the CCMV-DOTA-Gd was separated from free Gd^{3+} by running the reaction product over SEC using storage buffer as the eluent.

Characterization of the Modified CCMV Capsids

Characterization of both purified capsids was performed by SEC, dynamic light scattering (DLS), transmission electron microscopy (TEM), and LC/MS. Typical data from these characterization techniques is shown in Fig. 4. DLS analysis was performed on a ZetaPlus (Brookhaven Instruments). TEM (Leo 912AB) of negatively-stained samples (1% uranyl acetate) was performed.

Characterization GdDOTA Reaction Site in CCMV-DOTA-Gd

After removal of unbound Gd^{3+} from a solution of CCMV-DOTA-Gd, inductively-coupled plasma mass spectrometry (ICP-MS) analysis was used to determine the total concentration of Gd^{3+} bound to the viral capsid (7500; Agilent Technologies). Protein concentrations of the CCMV-

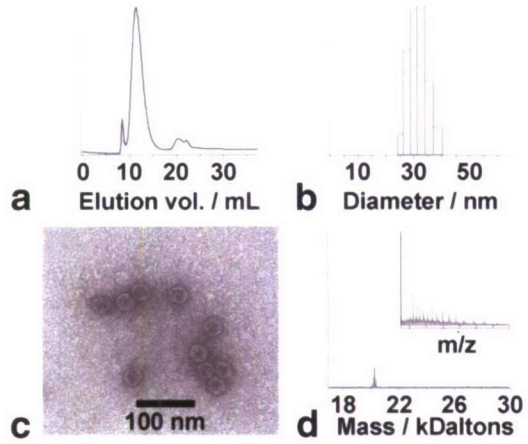


FIG. 4. Data from routine virus capsid characterization of CCMV-CAL or CCMV-DOTA. **a:** Size-exclusion chromatogram showing the 280-nm absorbance of the three main components. The small left peak is aggregated virus capsids. The large middle peak is intact capsids eluting at the correct retention volume and the peaks to the right are buffer molecules. **b:** DLS indicates a viral capsid mean diameter of 30 nm. **c:** Transmission electron micrograph of negatively-stained viral particles. **d:** The deconvoluted spectrum shows the correct subunit mass. Inset is the raw electrospray mass spectrum of viral capsid subunits.

DOTA-Gd samples were determined by the bicinchoninic acid (BCA) assay (Pierce). Protease digestion (Trypsin Gold; Promega) of CCMV-DOTA-Gd was carried out (1 mg/ml protein subunit, 1:200 subunit:trypsin, 37°C, 12 h) and liquid chromatography/tandem mass spectrometry LC/MS/MS was performed on the digested sample (nanoAquity coupled to a Q-Tof Premier). Data was analyzed with PLGS2 and MassLynx (Waters).

Generation of Binding Isotherms for CCMV-CAL

The lanthanides ions Gd^{3+} and terbium (Tb^{3+}) are known to bind with similar affinities; however, Gd^{3+} does not undergo FRET and therefore cannot be used to probe metal binding. Therefore Tb^{3+} was used as a Gd^{3+} mimic to study metal binding. All fluorescence experiments were performed on a (SPEX Fluorolog) spectrophotometer at 25°C. The fluorescence spectrum of CCMV-CAL-Tb was measured (λ max = 340 nm) following excitation at 295 nm, with a Tb^{3+} emission max was near 550 nm. The 340 nm peak was recorded by scanning from 305 nm to 575 nm in 1 nm steps. The 550 peak was monitored by scanning from 525 nm to 575 nm in 0.2 nm steps. Excitation and emission slit widths were set to 4 and 8 nm, respectively. Fluorescence was measured on 500 µl solutions of viral capsid to which metal ion was added in 5–20 µL increments from 10–500 µM solution standards. Subunit protein concentrations of 0.05 µM and 0.1 µM for the CCMV-CAL mutant were used.

Calculation of K_d

In analyzing the data for Tb^{3+} binding to the CCMV-CAL capsid, we have assumed that when metal ions are titrated to the CCMV-CAL capsid, they first bind completely to the inserted sites. Once the inserted sites are completely occupied then the endogenous sites start to bind the additional metal ions that are added. Equation [2] was used to analyze the data where θ is the fraction of capsid subunits with bound Tb^{3+}, $[Tb^{3+}{}_{Free}]$ is the free Tb^{3+} concentration, and K_d is the dissociation constant for subunit binding Tb^{3+} ions.

$$\theta = \frac{\left[Tb^{3+}_{Free}\right]}{[Tb^{3+}_{Free}] + K_d}. \qquad [1.2]$$

Baseline correction was performed on the 550 nm Tb^{3+} spectra. To find the maximum intensity of the 550-nm peak, a Gaussian function was individually fit to each spectra. The maximum intensity of the 550-nm peak (I_{550}) vs. the total Tb^{3+} in solution ($[Tb^{3+}{}_{total}]$) was plotted and fit to Eq. [3], where $K_{d_Initial}$ and $I_{550Limit}$ were the fitting parameters.

$$I_{550} = I_{550Limit}\left(\frac{[Tb^{3+}_{Total}]}{[Tb^{3+}_{Total}] + K_{d_Initial}}\right). \qquad [1.3]$$

Fractions bound terms (θ) were calculated by dividing I_{550} values by $I_{550Limit}$ determined from Eq. [3]. Values for $[Tb^{3+}{}_{Free}]$ were then calculated by Eq. [4].

$$[Tb^{3+}_{Free}] = [Tb^{3+}_{Total}] - (\theta[subunit_{Total}]). \qquad [1.4]$$

A plot of θ vs. $[Tb^{3+}{}_{free}]$ was then fit with the Eq. [2] and K_d was determined, along with an error associated with the fit.

STIOCHIOMETRIC TITRATION OF CCMV-CAL

Terbium ions (Tb^{3+}) were used as Gd^{3+} ion mimics. Tb^{3+} titrations were performed with the condition of [subunit] $>> K_d$ to determine the number of Tb^{3+} bound per CCMV-CAL subunit. Two titrations were performed in which the protein concentration used was 2.6 µM and 10 µM. The capsid was titrated to ~20 µM total Tb^{3+} in both experiments. The data, from both the beginning and end portions of the titration, were fit to linear functions.

RELAXOMETRY AND GD^{3+} QUANTITATION

For relaxometry experiments, fully assembled CCMV-CAL (60 µM subunit) containing 200 µM $GdCl_3$ was prepared in pH 6.5 buffer (50 mM HEPES, 150 mM NaCl). As a control, 200 µM $GdCl_3$ was prepared in the same buffer. CCMV-DOTA-Gd was prepared in pH 6.5 buffer (100 mM HEPES, 100 mM NaCl) with a subunit concentration of 101 µM (determined by the BCA assay) and 34 µM Gd^{3+} (determined by ICP-MS). Using a custom-designed variable field relaxometer, T_1 relaxivity was measured using a saturation-recovery pulse sequence with 32 incremental τ values. The range of Larmor frequencies was 2–62 MHz (0.05–1.5 T) and the measurements were carried out at a temperature of 23°C. T_1 values were determined by fitting data into Eq. [5] with A and B as fitting parameters.

$$f(Seconds) = A\left(1 - e\left(\frac{-Seconds}{\frac{1}{T1}}\right)\right) + B. \quad [1.5]$$

T_2 was measured using a CPMG pulse sequence with 500 echoes and an interecho time of 2 ms. T_2 values were determined by fitting data into Eq. [6] with A, B, and N as fitting parameters.

$$f(Seconds) = \sqrt{A\left(e\left(\frac{-2Seconds}{\frac{1}{T2}}\right) + N\right) + B} \quad [1.6]$$

Since the R1 and R2 relaxivities are expressed in units of (mM^{-1} of bound Gd^{3+} * s^{-1}), it was necessary to determine the mM concentration of bound Gd^{3+}. The calculation of the fraction of CCMV-CAL with bound Gd^{3+} turned into an approximation since this capsid contains two types of binding sites. First, the parameters for the higher affinity binding site or "CAL" were input into Eq. [7]. The concentration of binding sites [*BS*], [$Gd^{3+}{}_{Total}$], and the K_d are all values input into this equation.

$$\theta = \frac{([BS] + [Tb^{3+}_{Total}] + K_d - \sqrt{([BS] + [Tb^{3+}_{Total}] + K_d)^2 - (4[BS][Tb^{3+}_{Total}])}]}{2[BS]}. \quad [1.7]$$

From this, the fraction bound term for the CAL binding site (θ_{CAL}) could be determined and was further used (Eq. [8]) to calculate the [$Gd^{3+}{}_{Free}$] left after the CAL site was maximally bound.

$$Gd^{3+}_{Free} = Gd^{3+}_{Total} - ([CAL] \cdot \theta_{CAL}). \quad [1.8]$$

The fraction of Gd^{3+} bound to the endogenous site ($\theta_{Endogenous}$) was calculated by setting [$Gd^{3+}{}_{Total}$], which was input into Eq. [7], equal to [$Gd^{3+}{}_{Free}$], which was determined by Eq. [8]. K_d and [*BS*] values for the endogenous site were input into Eq. [7] to determine the fraction bound for the endogenous site ($\theta_{Endogenous}$). Finally, the fraction-bound terms for both the endogenous site ($\theta_{Endogenous}$) and the engineered site (θ_{CAL}) were multiplied by the millimolar concentration of their respective binding sites and then added together resulting in a millimolar concentration of total bound Gd^{3+}.

RESULTS

CCMV Modified Architecture

The CCMV viral capsid (Fig. 1) has been modified for enhanced Gd^{3+} binding, using two complementary approaches, while maintaining the advantages of the large molecular platform. In the first approach, a Gd^{3+} binding peptide from calmodulin, was genetically introduced onto the N-terminus of the CCMV viral capsid subunit (CCMV-CAL-Gd). This modified viral capsid has an increased affinity for Gd^{3+} in comparison with wild-type CCMV. In a second approach, GdDOTA was conjugated to CCMV resulting in high-affinity Gd^{3+} binding and imparting highly efficient relaxivity properties to the CCMV capsid (CCMV-DOTA-Gd).

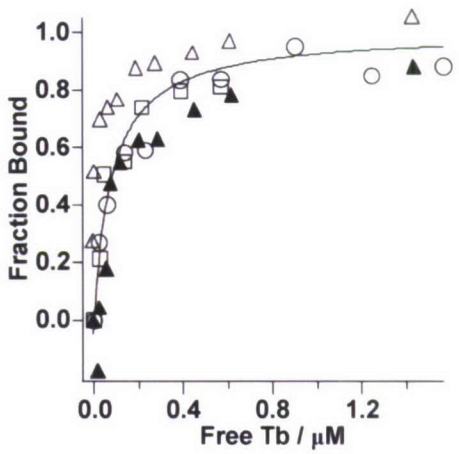


FIG. 5. Binding isotherms for CCMV-CAL binding Tb^{3+} ions. The open triangle, square, and circle represent three replicate experiments for the increase of the 550-nm peak. The solid triangle represents the decrease in the 340-nm peak from the first experiment. All four data sets were used for the fit.

Genetic Modification of CCMV—Attachment of Calmodulin Peptide

A peptide sequence from the Ca^{2+} binding portion of the protein calmodulin (DKDGDGWLEFEEGGG) was genetically fused to the N-terminus of CCMV (Fig. 2). Interestingly, this construct with a nine-residue peptide incorporated as an N-terminal fusion did not disrupt the ability of CCMV to self assemble as shown by SEC, DLS, and TEM analyses (Fig. 4). The mutation of the coat protein subunit gene was confirmed by DNA sequencing. LC/MS of the purified viral capsid subunit produced an experimental average mass of 20,234 Da compared to a calculated average mass of 20,232 Da (Supplemental Data) for the mutant protein subunit confirming the identity of the recombinant protein.

K_d Determination of CCMV-CAL

The metal binding affinity of the CCMV-CAL mutant was probed by FRET using excitation of endogenous tryptophan residues. The lanthanide ions Gd^{3+} and terbium (Tb^{3+}) are known to bind with similar affinities and both show preference for Ca^{2+} binding sites in proteins although with significantly higher affinities than Ca^{2+} binding (17,25). Titration of the CCMV-CAL mutant with increasing Tb^{3+} revealed a decrease in the tryptophan fluorescence (340 nm) and concomitant increase in the Tb^{3+} fluorescence (at 550 nm) indicating energy transfer between these sites (Fig. 5). The complete data set was fit to Eq. [2] and an average K_d of 82 ± 14 nM for Tb^{3+} binding to CCMV-CAL was determined. This indicates an enhancement in the metal binding affinity of 232-fold over binding to endogenous sites in the wild-type CCMV (K_d of 19,000 nM for Tb^{3+}) (21).

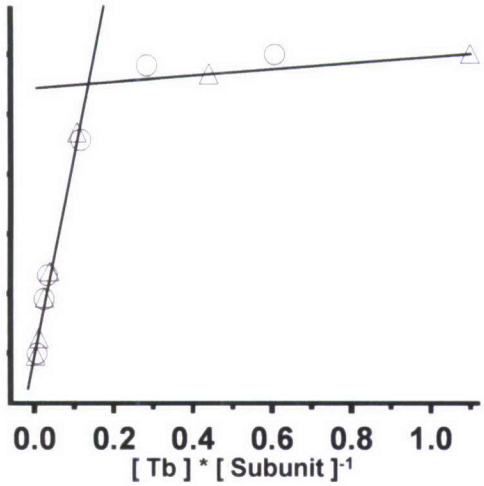


FIG. 6. A stoichiometric titration of CCMV CAL with Tb^{3+} ions. Two replicate data sets are shown here, one with circles and the other with triangles.

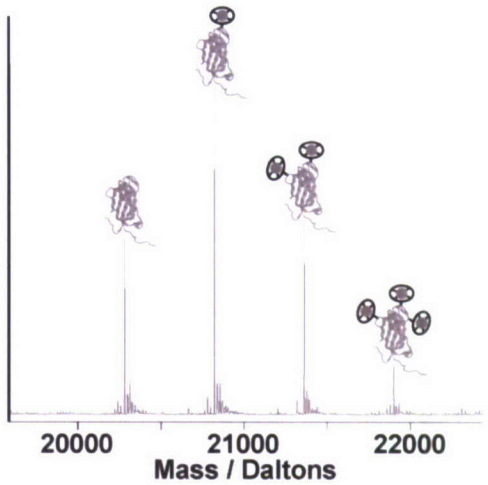


FIG. 7. Typical deconvoluted electrospray mass spectra of CCMV capsid subunits. CCMV capsids were reacted with NHS-DOTA and then $GdCl_3$. Unlabelled subunits and subunits with one to three GdDOTA were detected.

Stoichiometric Titrations of CCMV-CAL

A stoichiometric titration of Tb^{3+} was performed to determine the number of ions bound per CCMV-CAL subunit. When normalized fluorescence intensity was plotted against the ratio of subunit:Tb^{3+} ([subunit]/[Tb^{3+}_{Total}]) as shown in (Fig. 6), the data shows two distinct regions. The fluorescence response during the first part of the titration ([binding site] >> [metal]) increases linearly with Tb^{3+} addition. The second linear portion of the titration ([metal >> [binding site]) shows a relatively constant fluorescent response with a smaller slope. The *x*-intersection of the fits to each of these two regions indicates the point at which the binding sites are maximally occupied. Our data indicate a value of 7.3 ± 1.1 ([subunit]/[Tb^{3+}_{Total}]) for this intersection. At the end of the Tb^{3+} titration, there is one Tb^{3+} ion for every 7.3 ± 1.1 subunits or each CCMV capsid has approximately 25 metal ions bound at the introduced peptide sites.

Chemical Modification of CCMV—Attachment of GdDOTA

GdDOTA was covalently attached to the surface of CCMV via reaction with surface exposed lysine residues (Fig 3). A deconvoluted mass spectrum after a routine CCMV-DOTA-Gd synthesis and removal of unbound Gd^{3+}showed a distribution of 0 to 3 GdDOTA per subunit (Fig. 7). The addition of a single GdDOTA added an experimental mass of 544 Da to the subunit molecular weight, which corresponds well with the theoretical value of 543.7 Da per GdDOTA. It was qualitatively observed by LC/MS that up to two DOTA on average per subunit produced a stable, fully assembled capsid yielding up to 360 GdDOTA per capsid.

Using mass spectrometry we were able to map the lysines which are the sites of labeling. LC/MS/MS analysis indicates that GdDOTA is primarily attached through Lys 8 and Lys 45 (shown in Supplemental Data). The distribution of GdDOTA labeling on CCMV-DOTA-Gd suggests that there is not complete occupancy of either Lys 45 or Lys 8 within the capsid and that additional unidentified residues are also labeled. However, residues 8 and 45 are likely the most prevalently labeled lysines in CCMV. Figure 8 shows the inside view of the CCMV capsid with GdDOTA modeled onto residue 45. The position of lysine 45 in the structure suggests that the attached GdDOTA resides on the interior of the capsid. Residue 8 is not shown since the N-terminus is disordered in the X-ray crystal structure of CCMV and its position is therefore uncertain.

Relaxometry of CCMV-CAL-Gd and CCMV-DOTA-Gd

Highly efficient T_1 and T_2 relaxivity properties were observed in both capsids (CCMV-CAL-Gd and CCMV-DOTA-Gd). Table 1 summarizes R1 and R2 values for Gd^{3+}/virus systems, including: wild-type CCMV-Gd, CCMV-CAL-Gd,

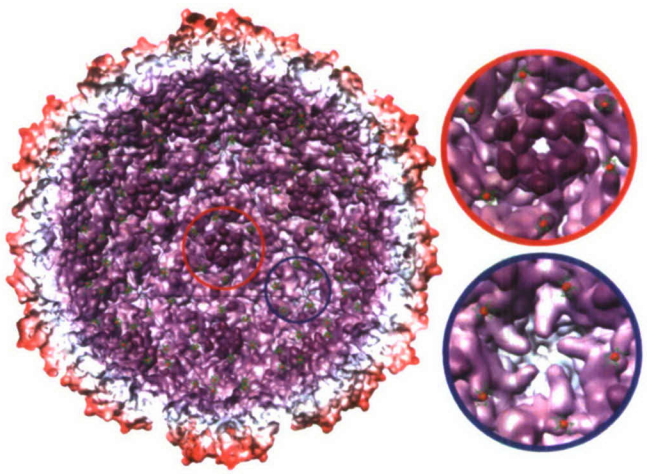


FIG. 8. Cutaway view of the interior of CCMV. GdDOTA was modeled to be attached through Lys 45. The red highlighted view is a close-up of Lys 45 in the six-fold environment while the blue highlighted view is of Lys 45 in a five-fold environment.

Table 1
Relaxivities of Capsid-Gd^{3+} Constructs at 1.5 T

Construct	Gd/particle[a]	r1 $(s^{-1})^*$ $(mM\ Gd^{3+})^{-1}$	r2 $(s^{-1})^*$ $(mM\ Gd^{3+})^{-1}$	r1 $(s^{-1})^*$ $(mM\ particle)^{-1}$	r2 $(s^{-1})^*$ $(mM\ particle)^{-1}$
Wild-type CCMV-Gd	141[a]	202	376	28,482	53,016
CCMV-CAL-Gd	172[a,b]	210	402[c]	36,120	69,144
CCMV-DOTA-Gd	61[a]	46	142	2806	8662
MS2-DTPA-Gd	514[a]	14	NA	7196	NA
DOTA-Gd	1	3[d]	5[d]	3[d]	5[d]

[a]Gd/particle (mole/mole) of the samples used in the relaxometry experiment.
[b]Calculated by setting CAL-sites/Gd = 7.3.
[c]Relaxivity measured at 1.3 T.
[d]Relaxivity measured at 1.0 T.
NA = Not available.

CCMV-DOTA-Gd, and MS2-DTPA-Gd (4,17). The ionic R1 and R2 trends are CCMV-CAL-Gd > CCMV-Gd > CCMV-DOTA-Gd > MS2-DTPA-Gd. The R1 and R2 trends relative to the particle are CCMV-CAL-Gd > wild-type CCMV-Gd > MS2-DTPA-Gd > CCMV-DOTA-Gd. In the genetic approach, the CCMV-CAL mutant exhibited approximately the same T_1 and T_2 relaxivity as wild-type CCMV-Gd previously reported (17). Figure 9a and b shows the similarities between the T_1 and T_2 relaxivity values of wild-type CCMV-Gd and CCMV-CAL-Gd for field strengths ranging from 0.2 to 1.5 T. This construct not only has an increased affinity for Gd^{3+}, but also maintains the very high relaxivity required for clinically relevant contrast agents.

The increased T_1 and T_2 relaxivity afforded by the chemical conjugation of GdDOTA to CCMV's capsid is shown (Fig. 9a and b). This conjugation resulted in increased T_1 and T_2 relaxivities by a factor of 10 relative to free GdDOTA. By chemically attaching GdDOTA to the CCMV capsid we have engineered a viral capsid with clinically relevant binding and high relaxivity.

The R1 dependence on field strength, for both wild-type CCMV-Gd and CCMV-CAL-Gd, varies in manner typical of nanoscale systems with a relaxivity maximum near 1 T (26). CCMV-DOTA-Gd has R1 field strength dependence more similar to small molecule systems with a maximum near 0.5 T. All CCMV/Gd systems exhibit a positive correlation between field strength and T_2 relaxivity values, as expected.

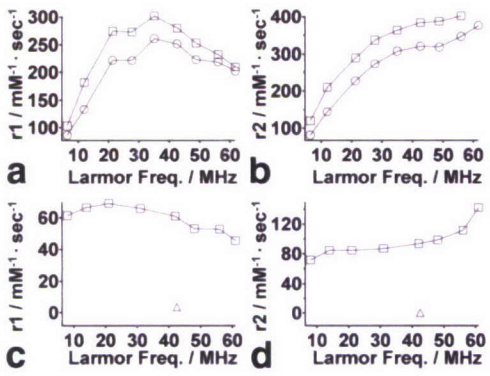


FIG. 9. **a:** R1 values for two CCMV capsids conjugated with Gd^{3+} ions, wild-type CCMV-Gd (circles) and CCMV-CAL-Gd (squares). **b:** R2 values for two CCMV capsids conjugated with Gd^{3+} ions, wild-type CCMV-Gd (circles), and CCMV-CAL-Gd (squares). **c:** R1 values for GdDOTA (triangle) and CCMV-DOTA-Gd (squares). **d:** R2 values for GdDOTA (triangle) and CCMV-DOTA-Gd (squares).

DISCUSSION

The major achievement of this study is demonstration that the CCMV viral capsid is a robust platform for development of MR contrast agents. Specifically we have shown that a genetically-engineered CCMV capsid binds Gd^{3+} (CCMV-CAL-Gd) with a significantly higher affinity than wild-type CCMV. In a second approach, the CCMV capsid has been chemically modified resulting in the attachment of multiple GdDOTA (CCMV-DOTA-Gd). These amine conjugated GdDOTA groups have been previously shown to bind with a dissociation constant of 10^{-20}, which corresponds to 10^{15} improvement in metal affinity compared to wild-type CCMV-Gd (27). Both these nanoparticle contrast agents are water soluble and have high ionic and particle relaxivities. These viral capsids have multiple Gd^{3+} ions attached per capsid, resulting in the potential for high contrast associated with tissue-specific localization. Localization of the viral capsid could be accomplished by active targeting or by passive means.

It is surprising that we observed substoichiometric binding of Gd^{3+} to the introduced CAL peptide in the CCMV-CAL-Gd construct. The data suggest that only 25 out of 180 sites bind Gd^{3+} at saturating conditions. A likely possibility is that not all the introduced sites are accessible to bind metal ions due to different chemical environments of the CAL peptides. It is also worth mentioning that the N-terminus is grouped in two environments, 20 pseudo six-fold and 12 five-fold environments (Fig. 2). The combinations of these groupings yields an average of 5.6 subunits per N-terminus grouping. This value is reasonably close to the experimentally determined value of 7.3 ± 1.1 CAL peptides per bound metal ion. It is possible that single metal ions are bound by multiple CAL peptides grouped at these five-fold and pseudo six-fold environments. The dissociation constant of 82 ± 14 nM for CCMV-CAL-Gd^{3+}, in comparison to dissociation constants in the range of 10^{-20} M for approved contrast agents, indicates that the CAL peptide binds Gd^{3+} too weakly for clinical application. Therefore, this approach of genetically attaching metal binding peptides to the CCMV capsid is not likely to be as fruitful as chemical modification approaches.

It was expected that all three CCMV/Gd constructs (CCMV-Gd, CCMV-CAL-Gd, and CCMV-DOTA-Gd) would have similar relaxivity values since size was thought to be the dominate factor in determining relaxivity rates. However, CCMV-DOTA-Gd has ionic relaxivity values that are approximately 25% of the values for CCMV-Gd and CCMV-CAL-Gd, indicating that factors other than size can influence the relaxivity of these capsids. The endogenous Gd^{3+} binding pocket in wild-type CCMV is at the interface of three subunits and contains side chains from each subunit. The result of this intersubunit binding pocket is that the overall motion of the Gd^{3+} ion is identical to the overall motion of the entire capsid and there is no additional motion of the Gd^{3+} ion. This is in contrast to the DOTA bound Gd^{3+} of CCMV-DOTA-Gd in which the flexibility of the linker connecting Gd to the capsid adds additional motion to the Gd^{3+} ion. In addition, LC/MS/MS data of a CCMV-DOTA-Gd trypsin digest indicates that the GdDOTA is primarily attached on the N-terminus end of the capsid subunit through Lys 8 and Lys 45. It is known that the N-terminus of CCMV is mobile and can occupy both the interior and exterior of the CCMV capsid architecture (28,29). This increased local mobility of the region of the viral capsid that is labeled with GdDOTA may cause a reduction in relaxivity. Electron paramagnetic resonance (EPR) spectroscopy studies of CCMV labeled with another small molecule concluded that there was local mobility within the spin label itself, which decreased the rotational correlation time when compared to the predicted value for the CCMV capsid (30–34). This local mobility of the label may also exist in CCMV-DOTA-Gd, which would lead to an additional decrease in relaxivity rates. Local mobility from flexible regions of the protein as well as the GdDOTA itself could explain the lower relaxivity values of CCMV-DOTA-Gd compared to CCMV-Gd and CCMV-CAL-Gd.

The consideration of gadolinium's ligands offers another explanation for the relaxivity differences between CCMV-DOTA-Gd and the two other CCMV/Gd^{3+} systems (CCMV-Gd and CCMV-CAL-Gd). Multiple water molecules are thought to be ligands of Gd^{3+} in both CCMV-Gd and CCMV-CAL-Gd in contrast to the GdDOTA system where a single water molecule is coordinated by the chelated Gd^{3+} atom. The increased number of bound water molecules makes CCMV-Gd and CCMV-CAL-Gd more efficient at relaxation of water molecules compared to CCMV-DOTA-Gd. Also, it was found that a peptide similar to the genetically fused sequence used in this work, resulted in unexpectedly high relaxivity values. These authors suggested that the high relaxivity values were a result of having the Gd^{3+} coordinated primarily with oxygen atoms since oxygen ligand systems resemble the aqua Gd^{3+} complex, which has near optimal lifetimes for bound water molecules (35–37). This ideal oxygen environment is found in CCMV-Gd and CCMV-CAL-Gd, whereas in the DOTA system one of the four carboxylate groups is converted to an amide bond in the conjugation reaction. This conversion increases the lifetime of the bound water molecule from approximately 250 ns to greater than 1 ms, which is further from the theoretical ideal lifetime of 20–30 ns (38). The larger value for the lifetime of bound water in the DOTA system probably accounts for some reduction of the relaxivity rate of CCMV-DOTA-Gd compared to CCMV-CAL-Gd.

Another possible example of how local mobility can affect relaxivity rates is revealed by comparing two different, Gd^{3+} chelated, viral capsid systems (CCMV-DOTA-Gd and MS2-DTPA-Gd). Anderson et al. (4) have attached up to 500 Gd^{3+} ions to the MS2 capsid yielding T_1 ionic relaxivity rates that are approximately three times lower than the T_1 ionic relaxivity rates for CCMV-DOTA-Gd (Table 1). This deviation is larger than expected since the relaxivity difference between DOTA and DTPA is small; the two capsids are the same size and in both systems the chelators are attached to endogenous lysines. Furthermore, the isothiocyanate-linked DTPA results in a lifetime for bound waters closer to the ideal value when compared to the NHS ester-linked DOTA system, so the MS2/DTPA system should result in higher relaxivity rates (26). A comparison of the length of the resulting linkers could explain the difference in observed relaxivity rates between the two cages. The NHS-ester used in the DOTA system results in a short linker with three rotatable bonds. This is in contrast to the DTPA system, in which the linker is longer and has four rotatable bonds. (The linker was measured from the lysine's amine nitrogen to the nitrogen that coordinates the Gd^{3+} ion on either chelator.) The longer and more rotatable linker in the DTPA system may result in more mobility and this could account for the higher ionic relaxivity rates for CCMV-DOTA-Gd compared to MS2-DTPA-Gd. Finally, Gd^{3+} loading of up to 360 Gd^{3+} ions per capsid has been achieved in the CCMV-DOTA-Gd construct, which yields a calculated particle relaxivity values of R1 = 16,560 and R2 = 51,120 $s^{-1}mM^{-1}$ at 1.5T.

Viral capsids offer advantages over other macromolecular contrast agents. Capsids are generally more rigid than liposome or dendrimer systems. These protein cages are homogeneous, whereas both dendrimers and liposomes are heterogeneous mixtures. Also the ability to accurately determine which residues are modified along with the availability of the near atomic resolution crystal structure provides information about the arrangement and environment of each modification within the viral capsid structure. This information aids in the design of the agent since the spatial arrangement, structural rigidity, and chemical environment of these modifications can be taken into account. In conclusion, the work presented here shows the potential of viral capsids as MRI contrast agents and will direct the design of the next generation of these imaging agents.

REFERENCES

1. Aime S, Cabella C, Colombatto S, Geninatti Crich S, Gianolio E, Maggioni F. Insights into the use of paramagnetic Gd(III) complexes in MR-molecular imaging investigations. J Magn Reson Imaging 2002;16:394–406.
2. Meade TJ, Taylor AK, Bull SR. New magnetic resonance contrast agents as biochemical reporters. Curr Opin Neurobiol 2003;13:597–602.
3. Yamada T, Iwasaki Y, Tada H, Iwabuki H, Chuah MK, VandenDriessche T, Fukuda H, Kondo A, Ueda M, Seno M, Tanizawa K, Kuroda S. Nanoparticles for the delivery of genes and drugs to human hepatocytes. Nat Biotechnol 2003;21:885–890.
4. Anderson EA, Isaacman S, Peabody DS, Wang EY, Canary JW, Kirshenbaum K. Viral nanoparticles donning a paramagnetic coat: Conjugation of MRI contrast agents to the MS2 capsid. Nano Lett 2006;6:1160–1164.

5. Lewis JD, Destito G, Zijlstra A, Gonzalez MJ, Quigley JP, Manchester M, Stuhlmann H. Viral nanoparticles as tools for intravital vascular imaging. Nat Med 2006;12:354–360.
6. Flenniken ML, Willits DA, Harmsen AL, Liepold LO, Harmsen AG, Young MJ, Douglas T. Melanoma and lymphocyte cell-specific targeting incorporated into a heat shock protein cage architecture. Chem Biol 2006;13:161–170.
7. Singh P, Destito G, Schneemann A, Manchester M. Canine parvovirus-like particles, a novel nanomaterial for tumor targeting. J Nanobiotechnology 2006;4:2.
8. Flenniken ML, Liepold LO, Crowley BE, Willits DA, Young MJ, Douglas T. Selective attachment and release of a chemotherapeutic agent from the interior of a protein cage architecture. Chem Commun (Camb) 2005:447–449.
9. Allen M, Willits D, Young M, Douglas T. Constrained synthesis of cobalt oxide nanomaterials in the 12-subunit protein cage from *Listeria innocua*. Inorg Chem 2003;42:6300–6305.
10. Douglas T, Young M. Viruses: making friends with old foes. Science 2006;312:873–875.
11. Flenniken ML, Willits DA, Brumfield S, Young MJ, Douglas T. The small heat shock protein cage from *Methanococcus jannaschii* is a versatile nanoscale platform for genetic and chemical modification. Nano Lett 2003;3:1573–1576.
12. Gillitzer E, Willits D, Young M, Douglas T. Chemical modification of a viral cage for multivalent presentation. Chem Commun (Camb) 2002: 2390–2391.
13. Klem MT, Willits D, Young M, Douglas T. 2-D array formation of genetically engineered viral cages on Au surfaces and imaging by atomic force microscopy. J Am Chem Soc 2003;125:10806–10807.
14. Raja KS, Wang Q, Gonzalez MJ, Manchester M, Johnson JE, Finn MG. Hybrid virus-polymer materials. 1. Synthesis and properties of PEG-decorated cowpea mosaic virus. Biomacromolecules 2003;4:472–476.
15. Sen Gupta S, Raja KS, Kaltgrad E, Strable E, Finn MG. Virus-glycopolymer conjugates by copper(I) catalysis of atom transfer radical polymerization and azide-alkyne cycloaddition. Chem Commun (Camb) 2005: 4315–4317.
16. Varpness Z, Peters JW, Young M, Douglas T. Biomimetic synthesis of a H-2 catalyst using a protein cage architecture. Nano Lett 2005;5:2306–2309.
17. Allen M, Bulte JWM, Liepold L, Basu G, Zywicke HA, Frank JA, Young M, Douglas T. Paramagnetic viral nanoparticles as potential high-relaxivity magnetic resonance contrast agents. Magn Reson Med 2005:54: 807–812.
18. Kobayashi H, Brechbiel MW. Dendrimer-based macromolecular MRI contrast agents: characteristics and application. Mol Imaging 2003;2:1–10.
19. Mulder WJ, Strijkers GJ, van Tilborg GA, Griffioen AW, Nicolay K. Lipid-based nanoparticles for contrast-enhanced MRI and molecular imaging. NMR Biomed 2006;19:142–164.
20. Speir JA, Munshi S, Wang G, Baker TS, Johnson JE. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. Structure 1995;3:63–78.
21. Basu G, Allen M, Willits D, Young M, Douglas T. Metal binding to cowpea chlorotic mottle virus using terbium(III) fluorescence. J Biol Inorg Chem 2003;8:721–725.
22. Rocklage SM, Worah D, Kim SH. Metal-ion release from paramagnetic chelates—what is tolerable. Magn Reson Med 1991;22:216–221.
23. Brumfield S, Willits D, Tang L, Johnson JE, Douglas T, Young M. Heterologous expression of the modified coat protein of Cowpea chlorotic mottle bromovirus results in the assembly of protein cages with altered architectures and function. J Gen Virol 2004;85:1049–1053.
24. Bancroft JB, Hiebert E, Rees MW, Markham R. Properties of cowpea chlorotic mottle virus its protein and nucleic acid. Virology 1968;34: 224–239.
25. Vazquez-Ibar JL, Weinglass AB, Kaback HR. Engineering a terbium-binding site into an integral membrane protein for luminescence energy transfer. Proc Natl Acad Sci USA 2002;99:3487–3492.
26. Laus S, Sour A, Ruloff R, Toth E, Merbach AE. Rotational dynamics account for pH-dependent relaxivities of PAMAM dendrimeric, Gd-based potential MRI contrast agents. Chemistry 2005;11:3064–3076.
27. Sherry AD, Brown RD, Geraldes CFGC, Koenig SH, Kuan K-T, Spiller M. Synthesis and characterization of the gadolinium(3+) complex of DOTA-propylamide: a model DOTA-protein conjugate. Inorg Chem 1989;28:620–622.
28. Liepold LO, Revis J, Allen M, Oltrogge L, Young M, Douglas T. Structural transitions in cowpea chlorotic mottle virus (CCMV). Phys Biol 2005;2:S166–S172.
29. Speir JA, Bothner B, Qu C, Willits DA, Young MJ, Johnson JE. Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics. J Virol 2006;80:3582–3591.
30. Vriend G, Hemminga MA, Haasnoot CAG, Hilbers CW. A two-dimensional nuclear Overhauser enhancement NMR-spectroscopy study at 500 MHz on cowpea chlorotic mottle virus protein assembled in spherical capsids. J Magn Reson 1985;64:501–505.
31. Vriend G, Hemminga MA, Verduin BJM, Schaafsma TJ. Swelling of cowpea chlorotic mottle virus studied by proton nuclear magnetic-resonance. FEBS Lett 1982;146:319–321.
32. Vriend G, Schilthuis JG, Verduin BJM, Hemminga MA. Saturation-transfer electron-spin-resonance spectroscopy on maleimide spin-labeled cowpea chlorotic mottle virus. J Magn Reson 1984;58:421–427.
33. Vriend G, Verduin BJM, Hemminga MA. Role of the N-terminal part of the coat protein in the assembly of cowpea chlorotic mottle virus—a 500 MHz proton nuclear-magnetic-resonance study and structural calculations. J Mol Biol 1986;191:453–460.
34. Hemminga MA, Faber AJ. Analysis of anisotropic spin-label motion in saturation-transfer electron-spin-resonance spectra of spin-labeled cowpea chlorotic mottle virus. J Magn Reson 1986;66:1–8.
35. Caravan P, Greenwood JM, Welch JT, Franklin SJ. Gadolinium-binding helix-turn-helix peptides: DNA-dependent MRI contrast agents. Chem Commun (Camb) 2003:2574–2575.
36. Lauffer RB. Paramagnetic metal-complexes as water proton relaxation agents for NMR imaging—theory and design. Chem Rev 1987;87:901–927.
37. Raymond KN, Pierre VC. Next generation, high relaxivity gadolinium MRI agents. Bioconjug Chem 2005;16:3–8.
38. Caravan P, Ellison JJ, McMurry TJ, Lauffer RB. Gadolinium(III) chelates as MRI contrast agents: structure, dynamics, and applications. Chem Rev 1999;99:2293–2352.
39. Congreve A, Parker D, Gianolio E, Botta M. Steric control of lanthanide hydration state: fast water exchange at gadolinium in a mono-amide 'DOTA' complex. Dalton Trans 2004:1441–1445.

Identification of the Dominant Precession-Damping Mechanism in Fe, Co, and Ni by First-Principles Calculations

K. Gilmore,[1,2] Y. U. Idzerda,[2] and M. D. Stiles[1]

[1]*National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8412, USA*
[2]*Physics Department, Montana State University, Bozeman, Montana 59717, USA*

(Received 4 May 2007; published 9 July 2007)

The Landau-Lifshitz equation reliably describes magnetization dynamics using a phenomenological treatment of damping. This Letter presents first-principles calculations of the damping parameters for Fe, Co, and Ni that quantitatively agree with existing ferromagnetic resonance measurements. This agreement establishes the dominant damping mechanism for these systems and takes a significant step toward predicting and tailoring the damping constants of new materials.

DOI: 10.1103/PhysRevLett.99.027204 PACS numbers: 75.40.Gb, 76.60.Es

Magnetic damping determines the performance of magnetic devices including hard drives, magnetic random access memories, magnetic logic devices, and magnetic field sensors. The behavior of these devices can be modeled using the Landau-Lifshitz (LL) equation [1]

$$\dot{\mathbf{m}} = -|\gamma|\mathbf{m} \times \mathbf{H}_{\text{eff}} - \frac{\lambda}{m^2}\mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{\text{eff}}), \quad (1)$$

or the essentially equivalent Gilbert (LLG) form [2,3]. The first term describes precession of the magnetization $\mathbf{m}$ about the effective field $\mathbf{H}_{\text{eff}}$, where $\gamma = g\mu_0\mu_B/\hbar$ is the gyromagnetic ratio. The second term is a phenomenological treatment of damping with the adjustable rate λ. The Gilbert form replaces this term with $\alpha\hat{m} \times \dot{\mathbf{m}}$ using the dimensionless damping constant $\alpha = \lambda/\gamma m$. The LL(G) equation adequately describes dynamics measured by techniques as varied as ferromagnetic resonance (FMR) [4], magneto-optical Kerr effect [5], x-ray absorption spectroscopy [6], and spin-current driven rotation with the addition of a spin-torque term [7,8].

Access to a range of damping rates in metallic materials is desirable when constructing devices for different applications. Empirically, doping NiFe alloys with transition metals [9] or rare earths [10] has produced compounds with damping rates in the range of $\alpha = 0.01$ to 0.8. A recent investigation of adding vanadium to iron resulted in an alloy with a decreased damping rate [11]. Unfortunately, the damping rate of a new material cannot be predicted because there has not yet been a first-principles calculation of damping that quantitatively agrees with experiment. The challenging pursuit of new materials with specific or lowered damping rates is further complicated by the expectation that, as device size continues to be scaled down, material parameters, such as λ, should change [12]. A detailed understanding of the important damping mechanisms in metallic ferromagnets and the ability to predictively calculate damping rates would greatly facilitate the design of new materials appropriate for a variety of applications.

The temperature dependence of damping in the transition metals has been carefully characterized through measurement of small angle dynamics by FMR [13,14]. While one might naïvely expect damping to increase monotonically with temperature, as it does for Fe, both Co and Ni also exhibit a dramatic rise in damping at low temperature as the temperature decreases. These observations indicate that two primary mechanisms are involved. Subsequent experiments [15,16] partition these nonmonotonic damping curves into a *conductivitylike* term that decreases with temperature like the conductivity and a *resistivitylike* term that increases with temperature like the resistivity. The two terms were found to give nearly equal weight to the damping curve of Ni.

A number of mechanisms for damping in these systems have been proposed [14,17–24]. See the review by Heinrich [25] for a more complete discussion. However, none of the models have been shown to quantitatively agree with measured values. The torque-correlation model of Kambersky [17] qualitatively matches the data, but has not been quantitatively evaluated in a rigorous fashion. Here, we report first-principles calculations of the Landau-Lifshitz damping constant according to Kambersky's torque-correlation expression. Quantitative comparison of the present calculations to the measured FMR values [13] positively identifies this damping pathway as the dominant effect in the transition metal systems. In addition to presenting these primary conclusions, we also describe the relationship between the torque-correlation model and the more widely understood breathing Fermi surface model [18,21], showing that the results of both models agree quantitatively in the low scattering rate limit.

The breathing Fermi surface model of Kambersky predicts

$$\lambda = \frac{g^2\mu_B^2}{\hbar}\sum_n \int \frac{dk^3}{(2\pi)^3}\eta(\epsilon_{n,k})\left(\frac{\partial\epsilon_{n,k}}{\partial\theta}\right)^2\frac{\tau}{\hbar}. \quad (2)$$

This model offers a qualitative explanation for the low temperature conductivitylike contribution to the measured

 © 2007 The American Physical Society

damping. The model describes damping of uniform precession as due to variations $\partial\epsilon_{n,k}/\partial\theta$ in the energies $\epsilon_{n,k}$ of the single-particle states with respect to the spin direction θ. The states are labeled with a wave vector k and band index n. As the magnetization precesses, the spin-orbit interaction changes the energy of the electronic states, pushing some occupied states above the Fermi level and some unoccupied states below the Fermi level. Thus, electron-hole pairs are generated near the Fermi level even in the absence of changes in the electronic populations. The η function in Eq. (2) is the negative derivative of the Fermi function and picks out only states near the Fermi level to contribute to the damping. g is the Landé g factor and μ_B is the Bohr magneton. The electron-hole pairs created by the precession exist for some lifetime τ before relaxing through lattice scattering. The amount of energy and angular momentum dissipated to the lattice depends on how far from equilibrium the system gets; thus, damping by this mechanism increases linearly with the electron lifetime as seen in Eq. (2). Since the electron lifetime is expected to decrease as the temperature increases, this model predicts that damping diminishes as the temperature is raised.

Because the predicted damping rate is linear in the scattering time, the damping rate cannot be calculated more accurately than the scattering time is known. For this reason it is not possible to make quantitative comparisons between calculations of the breathing Fermi surface and measurements. Further, while the breathing Fermi surface model can explain the dramatic temperature dependence observed in the conductivitylike portion of the data it fails to capture the physics driving the resistivitylike term. This is a significant limitation from a practical perspective because the resistivitylike term dominates damping at room temperature and above and is the only contribution observed in iron [13] and NiFe alloys [26].

Kambersky's torque-correlation model predicts

$$\lambda = \frac{g^2\mu_B^2}{\hbar}\sum_{n,m}\int\frac{dk^3}{(2\pi)^3}|\Gamma^-_{nm}(k)|^2 W_{nm}(k) \qquad (3)$$

and we will show that it both incorporates the physics of the breathing Fermi surface model and also accounts for the resistivitylike terms. The matrix elements $\Gamma^-_{nm}(k) = \langle n,k|[\sigma^-, H_{so}]|m,k\rangle$ measure transitions between states in bands n and m induced by the spin-orbit torque. These transitions conserve wave vector k because they describe the annihilation of a uniform precession magnon, which carries no linear momentum. The nature of these scattering events, which are weighted by the spectral overlap $W_{nm}(k) = (1/\pi)\int d\omega_1 \eta(\omega_1)A_{nk}(\omega_1)A_{mk}(\omega_1)$, will be discussed in more detail below. The electron spectral functions A_{nk} are Lorentzians centered around the band energies ϵ_{nk} and broadened by interactions with the lattice. The width of the spectral function $\hbar/\tau$ provides a phenomenological account for the role of electron-lattice scattering in the damping process. The η function is the same as in Eq. (2) and enforces the requirement of spectral overlap at the Fermi level.

Equation (3) captures two different types of scattering events: scattering within a single band, $m = n$, for which the initial and final states are the same, and scattering between two different bands, $m \neq n$. As explained in [17] the overlap of the spectral functions is proportional (inverse) to the electron scattering time for intraband (interband) scattering. From this observation the qualitative conclusion is made that the intraband contributions match the conductivitylike terms while the interband contributions give the resistivitylike terms. Evaluation of Eq. (3) is more computationally intensive than that of the breathing Fermi surface model and until now only a few estimates for Ni and Fe have been made [19].

We have performed first-principles calculations of the torque-correlation model Eq. (3) with realistic band structures for Fe, Co, and Ni. Prior to evaluating Eq. (3) the eigenstates and energies of each metal were found using the linear augmented plane wave method [27] in the local spin density approximation [28–30]. Details of the calculations for these materials are described in [31]. The exchange field was fixed in the chosen equilibrium magnetization direction. Calculations of Eq. (3) presented in this Letter are converged to within a standard deviation of 3%, which required sampling $(160)^3$ k points for Fe, $(120)^3$ for Ni, and $(100)^2$ k points in the basal plane by 57 along the c axis for Co. Electron-lattice interactions were treated phenomenologically as a broadening of the spectral functions. The Fermi distribution was smeared with an artificial temperature. Results did not vary significantly with reasonable choices of this temperature since the broadening of the Fermi distribution was considerably less than that of the bands. The damping rate was calculated for a range of scattering rates (spectral widths) just as damping has been measured over a range of temperatures.

The results of these calculations are presented in Fig. 1 and are decomposed into the intraband and interband terms. The downward sloping line in Fig. 1 represents the intraband contribution to damping. Damping constants were recently calculated using the breathing Fermi surface model [12,21] by evaluating the derivative of the electronic energy with respect to the spin direction according to Eq. (2). The results of the breathing Fermi surface prediction are indistinguishable from the intraband terms of the present calculation even though the computational approaches differed significantly; the agreement is quantified in Table I.

The breathing Fermi surface model could not be quantitatively compared to the experimental results because the temperature dependence of the scattering rate has not been determined sufficiently accurately. While the present calculations also require knowledge of the scattering rate to determine the damping rate, the nonmonotonic dependence of damping on the scattering rate produces a unique minimum damping rate. In the same manner that the calculated curves of Fig. 1 have a minimum with respect

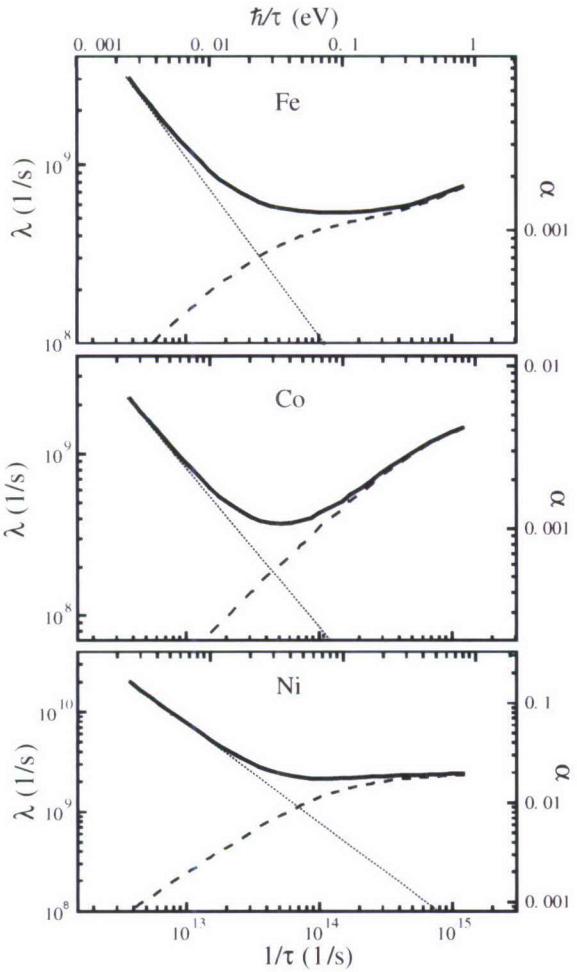


FIG. 1. Calculated Landau-Lifshitz damping constant for Fe, Co, and Ni. Thick solid curves give the total damping parameter while dotted curves give the intraband and dashed lines the interband contributions. Values for λ are given in SI units. The right axis is the equivalent Gilbert damping parameter and the top axis is the full width half maximum of the electron spectral functions.

to scattering rate, the measured damping curves exhibit minima with respect to temperature. Whatever the relation between temperature and scattering rate, the calculated minima may be compared directly and quantitatively to the measured minima. Table I makes this comparison. The agreement between measured and calculated values shows that the torque-correlation model accounts for the dominant contribution to damping in these systems.

Our calculated values are smaller than the measured values. Using measured g values instead of setting $g = 2$ would increase our results by a factor of $(g/2)^2$, or about 10% for Fe and 20% for Co and Ni. Other possible reasons for the difference include a simplified treatment of electron-lattice scattering in which the scattering rates for all states were assumed equal, errors associated with using wave functions found from the mean-field local spin density approximation, and numerical convergence (discussed above). Additionally, the extraction of damping rates from the measured linewidths remains challenging for Fe and particularly for Co. Other damping mechanisms may also make small contributions [22–24].

Since the manipulations involved with the equation of motion techniques employed in deriving Eq. (3) obscure the underlying physics we now discuss the two scattering processes and connect the intraband terms to the breathing Fermi surface model. The intraband terms in Eq. (3) describe scattering from one state to itself by the torque operator, which is similar to a spin-flip operator. A spin-flip operation between some state and itself is only nonzero because the spin-orbit interaction mixes small amounts of the opposite spin direction into each state. Since the initial and final states are the same, the operation is naturally spin conserving. The matrix elements do not describe a real transition, but rather provide a measure of the energy of the electron-hole pairs that are generated as the spin direction changes. The electron-hole pairs are subsequently annihilated by a real electron-lattice scattering event.

To connect the derivatives $\partial\epsilon/\partial\theta$ in Eq. (2) and the torque matrix elements in Eq. (3) we imagine first pointing the magnetization in some direction $\hat{z}$. The only energy that changes with the magnetization direction is the spin-orbit energy H_{so}. As the spin of a single-particle state $|\rangle$ rotates along $\hat{\theta}$ about $\hat{x}$ its spin-orbit energy is given by $\epsilon(\theta) = \langle|e^{i\sigma_x\theta}H_{\mathrm{so}}e^{-i\sigma_x\theta}|\rangle$. The derivative with respect to θ is $\partial\epsilon(\theta)/\partial\theta = i\langle|e^{i\sigma_x\theta}[\sigma_x, H_{\mathrm{so}}]e^{-i\sigma_x\theta}|\rangle$. Evaluating this derivative at the pole ($\theta = 0$) gives $\partial\epsilon/\partial\theta = i\langle|[\sigma_x, H_{\mathrm{so}}]|\rangle$. Similarly, rotating the spin along $\hat{\theta}$ about $\hat{y}$ leads to

TABLE I. Calculated and measured [13] damping parameters. Values for λ are reported in 10^9 s^{-1} while those for α are dimensionless. Values in the first four columns indicate minima of the calculated or measured curves. The last two columns list calculated damping due to the intraband contribution from Eq. (3) and from the breathing Fermi surface model [12], respectively. Values for λ/τ are given in 10^{22} s^{-2}. Published numbers from [12,13] have been multiplied by 4π to convert from the cgs unit system to SI.

	α_{calc}	λ_{calc}	λ_{meas}	$\lambda_{\mathrm{calc}}/\lambda_{\mathrm{meas}}$	$(\lambda/\tau)_{\mathrm{intra}}$	$(\lambda/\tau)_{\mathrm{BFS}}$
bcc Fe ⟨001⟩	0.0013	0.54	0.88	0.61	1.01	0.968
bcc Fe ⟨111⟩	0.0013	0.54	⋯	⋯	1.35	1.29
hcp Co ⟨0001⟩	0.0011	0.37	0.9	0.41	0.786	0.704
fcc Ni ⟨111⟩	0.017	2.1	2.9	0.72	6.67	6.66
fcc Ni ⟨001⟩	0.018	2.2	⋯	⋯	8.61	8.42

$\partial\epsilon/\partial\theta = i\langle|[\sigma_y, H_{so}]|\rangle$. The torque matrix elements in Eq. (3) are $\Gamma^- = \langle|[\sigma^-, H_{so}]|\rangle = \langle|[\sigma_x, H_{so}]|\rangle - i\langle|[\sigma_y, H_{so}]|\rangle$. Using the relations between the commutators and derivatives just found the torque is $\Gamma^- = -i(\partial\epsilon/\partial\theta)_x - (\partial\epsilon/\partial\theta)_y$, where the subscripts indicate the rotation axis. Squaring the torque matrix elements gives $|\Gamma^-|^2 = (\partial\epsilon/\partial\theta)_x^2 + (\partial\epsilon/\partial\theta)_y^2$. For high symmetry directions $(\partial\epsilon/\partial\theta)_x = (\partial\epsilon/\partial\theta)_y$ and we deduce $|\Gamma^-|^2 = 2(\partial\epsilon/\partial\theta)^2$ demonstrating that the intraband terms of the torque-correlation model describe the same physics as the breathing Fermi surface.

The monotonically increasing curves in Fig. 1 indicate the interband contribution to damping. Uniform mode magnons, which have negligible energy, may induce quasielastic transitions between states with different energies. This occurs when lattice scattering broadens bands sufficiently so that they overlap at the Fermi level. These wave vector conserving transitions, which are driven by the precessing exchange field, occur primarily between states with significantly different spin character. The process may roughly be thought of as the decay of a uniform precession magnon into a single electron spin-flip excitation. These events occur more frequently as the band overlaps increase. For this reason the interband terms, which qualitatively match the resistivitylike contributions in the experimental data, dominate damping at room temperature and above.

We have calculated the Landau-Lifshitz damping parameter for the itinerant ferromagnets Fe, Co, and Ni as a function of the electron-lattice scattering rate. The intraband and interband components match qualitatively to conductivity and resistivitylike terms observed in FMR measurements. A quantitative comparison was made between the minimal damping rates calculated as a function of scattering rate and measured with respect to temperature. This comparison demonstrates that our calculations account for the dominant contribution to damping in these systems and identify the primary damping mechanism. At room temperature and above, damping occurs overwhelmingly through the interband transitions. The contribution of these terms depends in part on the band gap spectrum around the Fermi level, which could be adjusted through doping.

K. G. and Y. U. I. acknowledge the support of the Office of Naval Research through Grant No. N00014-03-1-0692 and through Grant No. N00014-06-1-1016. We would like to thank R. D. McMichael and T. J. Silva for valuable discussions.

[1] L. Landau and E. Lifshitz, Phys. Z. Sowjetunion **8**, 153 (1935).
[2] T. L. Gilbert, Armour Research Foundation Project No. A059, 1956 (unpublished).
[3] T. L. Gilbert, IEEE Trans. Magn. **40**, 3443 (2004).
[4] D. J. Twisselmann and R. D. McMichael, J. Appl. Phys. **93**, 6903 (2003).
[5] T. Gerrits, J. Hohlfeld, O. Gielkens, K. J. Veenstra, K. Bal, T. Rasing, and H. A. M. van den Berg, J. Appl. Phys. **89**, 7648 (2001).
[6] W. E. Bailey, L. Cheng, D. J. Keavney, C. C. Kao, E. Vescovo, and D. A. Arena, Phys. Rev. B **70**, 172403 (2004).
[7] I. N. Krivorotov, D. V. Berkov, N. L. Gorn, N. C. Emley, J. C. Sankey, D. C. Ralph, and R. A. Buhrman, arXiv:cond-mat/0703458 [Phys. Rev. B (to be published)].
[8] M. D. Stiles and J. Miltat, *Spin Dynamics in Confined Magnetic Structures III* (Springer, New York, 2006).
[9] J. O. Rantschler, R. D. McMichael, A. Castiello, A. J. Shapiro, W. F. Egelhoff, B. B. Maranville, D. Pulugurtha, A. P. Chen, and L. M. Conners, J. Appl. Phys. **101**, 033911 (2007).
[10] W. Bailey, P. Kabos, F. Mancoff, and S. Russek, IEEE Trans. Magn. **37**, 1749 (2001).
[11] C. Scheck, L. Cheng, I. Barsukov, Z. Frait, and W. E. Bailey, Phys. Rev. Lett. **98**, 117601 (2007).
[12] D. Steiauf and M. Fähnle, Phys. Rev. B **72**, 064450 (2005).
[13] S. M. Bhagat and P. Lubitz, Phys. Rev. B **10**, 179 (1974).
[14] B. Heinrich, D. Fraitová, and V. Kamberský, Phys. Status Solidi **23**, 501 (1967).
[15] B. Heinrich, D. J. Meredith, and J. F. Cochran, J. Appl. Phys. **50**, 7726 (1979).
[16] J. F. Cochran and B. Heinrich, IEEE Trans. Magn. **16**, 660 (1980).
[17] V. Kamberský, Czechoslovak Journal of Physics, Section B **26**, 1366 (1976).
[18] V. Kamberský, Can. J. Phys. **48**, 2906 (1970).
[19] V. Kamberský, Czechoslovak Journal of Physics, Section B **34**, 1111 (1984).
[20] V. Korenman and R. E. Prange, Phys. Rev. B **6**, 2769 (1972).
[21] J. Kuneš and V. Kamberský, Phys. Rev. B **65**, 212411 (2002).
[22] R. D. McMichael and A. Kunz, J. Appl. Phys. **91**, 8650 (2002).
[23] E. Rossi, O. G. Heinonen, and A. H. MacDonald, Phys. Rev. B **72**, 174412 (2005).
[24] Y. Tserkovnyak, G. A. Fiete, and B. I. Halperin, Appl. Phys. Lett. **84**, 5234 (2004).
[25] B. Heinrich, *Ultrathin Magnetic Structures III* (Springer, New York, 2005).
[26] S. Ingvarsson, L. Ritchie, X. Y. Liu, G. Xiao, J. C. Slonczewski, P. L. Trouilloud, and R. H. Koch, Phys. Rev. B **66**, 214416 (2002).
[27] L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **33**, 823 (1986).
[28] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[29] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[30] U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).
[31] M. D. Stiles, S. V. Halilov, R. A. Hyman, and A. Zangwill, Phys. Rev. B **64**, 104430 (2001).

Efficient Immobilization and Patterning of Live Bacterial Cells

Zhiyong Suo,[†] Recep Avci,*[,†] Xinghong Yang,[‡] and David W. Pascual[‡]

Imaging and Chemical Analysis Laboratory, Department of Physics, and Veterinary Molecular Biology, Montana State University, Bozeman, Montana 59717

Received December 31, 2007. In Final Form: January 28, 2008

A monolayer of live bacterial cells has been patterned onto substrates through the interaction between CFA/I fimbriae and the corresponding antibody. Patterns of live bacteria have been prepared with cellular resolution on silicon and gold substrates for *Salmonella enterica* serovar Typhimurium as a model with high specificity and efficiency. The immobilized cells are capable of dividing in growth medium to form a self-sustaining bacterial monolayer on the patterned areas. Interestingly, the immobilized cells can alter their orientation on the substrate, from lying-down to standing-up, as a response to the cell density increase during incubation. This method was successfully used to sort a targeted bacterial species from a mixed culture within 2 h.

Introduction

Bacterial cells are ideal sensors for environmental monitoring because of their low cost, fast growth, rich genetic modifications, easy handling, and sensitivity to a wide variety of environmental stimuli.[1] Efficient, controllable immobilization of bacteria is critical for the success of such biosensors. Such immobilization also offers potential applications in biomedical research and fundamental bacteriological studies such as quorum sensing. The majority of reported immobilization approaches utilize either nonspecific adsorption of bacterial cells on chemically treated surfaces or physical entrapment of cells in gels or microholes. For example, attachment of bacteria has been conducted on prefabricated microarrays with microholes treated either with poly-L-lysine (PLL)[2] or with *n*-hexadecanethiol,[3] while the areas surrounding these holes were coated with a layer of poly(ethylene glycol) (PEG). Micro-[4] and macrocontact[5] printing have been employed to transfer live bacteria onto the surface of a nutrient-containing matrix such as agarose or hydrogel. Bacterial microarrays have also been prepared by loading individual bacterial cells into microwells (2.5 μm wide, ~3 μm deep) at the distal end of an optical fiber bundle by centrifugation.[6] A bacterial array printed onto porous nylon has also been reported,[7] in which cells were physically entrapped in the pores of a special nylon substrate in close contact with a nutrient medium. Such arrays offer great potential for monitoring genotoxicity[8] and heavy metals in the environment[9] and for high throughput assays of gene expression.[10] Very recently, Akselrod et al. reported three-dimensional heterotypic arrays of living cells in hydrogels created by means of high-precision (submicrometer accuracy) time-multiplexed holographic laser trapping.[11] However, this technique has limited applications in practice, in that, besides the need for a trapping laser, excessive exposure to laser light may cause photodamage to the cells; furthermore, arrays are expected to merge in a few hours because of cell division.

Another approach to bacterial immobilization takes advantage of the interaction between an appropriate antibody−antigen pair. This approach has been used in conjunction with microcontact printing and dip-pen nanolithography;[12] binding of *Escherichia coli* to microscale features was achieved through antibodies against the whole cell or bacterial flagella. However, *E. coli* cells showed a lower attachment to features modified with antibodies than with PLL.[12] The poor immobilization of bacterial cells mediated by antibody binding is also evidenced by a report on environmental toxicity monitoring using immobilized *E. coli* in which only 2% surface coverage of the bacteria was achieved.[13] In other studies, antibody-modified substrates have been used for immobilizing and detecting pathogenic bacteria, but little has been reported on the cell density on the substrates.[14] The present study, as well as previous studies, indicates that antibody−antigen-based immobilization does not hinder such physiological activities as cell division (this work), gene expression, or bioluminescence of bacteria at the locations of their immobilization.[12,13]

In most of these applications, limited attention has been paid to the efficiency and control of the immobilization of *live* cells on substrates or to the physiological activity of individually immobilized bacteria. For example, patches of bacteria created by microcontact printing will either grow in lateral directions or disintegrate in a short period of time if exposed to a flow reactor, resulting in a loss of functionality of the sensor. An efficient, reproducible, stable, self-sustaining, and highly specific immobilization of bacteria on a predefined surface is necessary. In this paper, we report such an immobilization of live cells of *S.*

* To whom correspondence should be addressed. E-mail: avci@physics.montana.edu. Telephone: 406-994-6164. Fax: 406-994-6040.

[†] Department of Physics.

[‡] Veterinary Molecular Biology.

(1) Belkin, S. *Curr. Opin. Microbiol.* **2003**, *6* (3), 206−212.

(2) Rozhok, S.; Fan, Z. F.; Nyamjav, D.; Liu, C.; Mirkin, C. A.; Holz, R. C. *Langmuir* **2006**, *22* (26), 11251−11254.

(3) Rowan, B.; Wheeler, M. A.; Crooks, R. M. *Langmuir* **2002**, *18* (25), 9914−9917.

(4) Xu, L.; Robert, L.; Ouyang, Q.; Taddei, F.; Chen, Y.; Lindner, A. B.; Baigl, D. *Nano Lett.* **2007**, *7* (7), 2068−2072.

(5) Weibel, D. B.; Lee, A.; Mayer, M.; Brady, S. F.; Bruzewicz, D.; Yang, J.; DiLuzio, W. R.; Clardy, J.; Whitesides, G. M. *Langmuir* **2005**, *21* (14), 6436−6442.

(6) Brogan, K. L.; Walt, D. R. *Curr. Opin. Chem. Biol.* **2005**, *9* (5), 494−500.

(7) Heitkamp, M. A.; Stewart, W. P. *Appl. Environ. Microbiol.* **1996**, *62* (12), 4659−4662.

(8) Kuang, Y.; Biran, I.; Walt, D. R. *Anal. Chem.* **2004**, *76* (10), 2902−2909.

(9) Biran, I.; Rissin, D. M.; Ron, E. Z.; Walt, D. R. *Anal. Biochem.* **2003**, *315* (1), 106−113.

(10) Van Dyk, T. K.; DeRose, E. J.; Gonye, G. E. *J. Bacteriol.* **2001**, *183* (19), 5496−5505.

(11) Akselrod, G. M.; Timp, W.; Mirsaidov, U.; Zhao, Q.; Li, C.; Timp, R.; Timp, K.; Matsudaira, P.; Timp, G. *Biophys. J.* **2006**, *91* (9), 3465−3473.

(12) Rozhok, S.; Shen, C. K. F.; Littler, P. L. H.; Fan, Z. F.; Liu, C.; Mirkin, C. A.; Holz, R. C. *Small* **2005**, *1* (4), 445−451.

(13) Premkumar, J. R.; Lev, O.; Marks, R. S.; Polyak, B.; Rosen, R.; Belkin, S. *Talanta* **2001**, *55* (5), 1029−1038.

(14) Oh, B. K.; Kim, Y. K.; Park, K. W.; Lee, W. H.; Choi, J. W. *Biosens. Bioelectron.* **2004**, *19* (11), 1497−1504.

10.1021/la7038653 CCC: $40.75 © 2008 American Chemical Society
Published on Web 03/06/2008

Typhimurium on well-characterized material surfaces; we chose this species because of its zoonotic properties, infecting both animals and humans, and our desire to prevent such infections.

Experimental Section

Bacteria. In most experiments, *Salmonella enterica* serovar Typhimurium Δ*asd*::*kan*^{R} H71-pHC was used as a model bacterial species for immobilization and patterning. We constructed a Δ*asd*::*kan*^{R} lethal mutant from wild type *S.* Typhimurium H71, termed Δ*asd*::*kan*^{R} *S.* Typhimurium H71.[15] This lethal mutant cannot normally survive unless an *asd* gene is presented *in trans*. Plasmid pHC, which contains a chimeric triple promoter, P*tetA*~P*pagC*~P*phoP*,[16,17] was used to express CFA/I fimbriae. The chimeric triple promoter was installed upstream of the *cfa*/I operon to enhance *cfa*/I expression. For the sorting experiment (Figure 6), *S.* Typhimurium Δ*asd*::*kan*^{R} H71-(pHC+pGFP) and *E. coli* O157:H7 RFP were used. Plasmids pHC and pGFP (Clontech, Mountain View, CA) were used to express CFA/I fimbriae and green fluorescence protein (GFP), respectively. *E. coli* O157:H7 RFP expressing red fluorescence protein (RFP)[18] was obtained from Dr. T. Khan and Dr. B. Klayman at the Center for Biofilm Engineering, Montana State University.

Frozen bacteria stock at −80 °C was inoculated onto a Luria−Bertani (LB) plate and incubated at 37 °C overnight. The bacteria were then inoculated into an LB liquid medium without antibiotics and shaken at 125 rpm at 37 °C. The bacterial cells were harvested when the optical density of the medium at 600 nm (OD600) reached about 0.5−0.6, which corresponds to a colony forming unit (CFU) value of ~9.0×10^8/mL.

Antibody. The anti-CFA/I serum was prepared by immunizing a rabbit intramuscularly (im) with purified CFA/I fimbriae proteins. Four weeks post immunization, the rabbit was bled to check the serum anti-CFA/I titers using an enzyme-linked immunoadsorbent assay (ELISA). Serum IgG was further purified with the protein G column to remove the nonspecific serum protein. This antibody was diluted to 100 times with phosphate-buffered saline (PBS) (pH = 7.4) before use.

Chemicals. PBS buffer salt, 3-aminopropyltriethoxysilane (APTES), and 11-mercaptoundecanoic acid (11-MUDA) were purchased from Sigma-Aldrich (St. Louis, MO). *N*-[β-Maleimidopropyloxy]-succinimide ester (BMPS), 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC), and *N*-hydroxysuccinimide (NHS) were purchased from Pierce Biotechnology (Rockford, IL). Bacterial viability dyes (Live/Dead *Bac*Light) were purchased from Invitrogen (Carlsbad, CA). $HS(CH_2)_{11}(OCH_2CH_2)_3$-OH (SPT-11) was purchased from Sensopath Technology, Inc. (Bozeman, MT), and 2-[methoxy(polyethyleneoxy)-propyl]trimethoxysilane (PEG-silane) was purchased from Gelest, Inc. (Morrisville, PA).

Substrate Passivation. *Gold.* Gold-coated silicon chips were cleaned in a chloroform bath and ozone plasma chamber, rinsed with 100% ethanol, and then incubated in a 2 mM solution of SPT-11[19] in 100% ethanol overnight. After being rinsed with copious (5 × 1 mL) ethanol and water, the modified chips were dried with nitrogen.

Silicon. Silicon chips cleaned by ozone plasma were heated in a solution of PEG-silane[20] (2%, v/v) in isopropanol at 60 °C for 70 min, rinsed with copious (5 × 1 mL) 100% ethanol, and dried with nitrogen.

Patterning Substrates Using a Microfocused Ga^+ Ion Beam. The substrate was patterned using a Ga^+ beam of a time-of-flight secondary ion mass spectrometry (ToFSIMS) system (TRIFT I, PHI-Evans, Chanhassen, MN).[21] The microfocused Ga^+ beam provides a 1.3 nA DC current at ~15 keV beam energy. The sample potential for positive ions was kept at ~3 kV; hence, the primary ion impact energy was ~12 keV. The direct Ga^+ ion current at the target position was measured to be ~1.26 nA. The etching time was carefully adjusted so that only a very thin layer of the substrate surface was removed (<5 nm). To the best of our knowledge this is the first report on the preparation of bacterial cell patterns using a focused ion beam.

Covalent Linking of Antibody. *Gold.* Etched gold substrates were incubated in a 2 mM solution of 11-MUDA in 100% ethanol overnight. After being rinsed with copious (5 × 1 mL) ethanol and water, the modified chips were treated with a NHS/EDC solution (NHS, 3 mg/mL; EDC, 2 mg/mL) for 1 h at ambient temperature and further rinsed with PBS buffer.

Silicon. Etched silicon substrates were incubated in a solution of APTES in methanol (2%) for 10 min and rinsed with copious (5 × 1 mL) ethanol before being further incubated in a solution of BMPS in anhydrous acetonitrile (10 mM) for 30 min at room temperature. The chips were rinsed with acetonitrile and dried in air.

Antibody Linking. The activated substrates were incubated with antibody solution for 1 h at room temperature and rinsed with PBS buffer to remove the free antibody molecules. These antibody-modified chips were then used for bacterial immobilization.

Immobilization of Bacterial Cells. Substrates with antibody patterns were incubated with a suspension of live bacterial cells in half LB growth medium (LB medium diluted with 1% NaCl solution by a ratio of 1:1, v/v) at room temperature for 3 h. Some samples were also incubated at 37 °C for 3 h and showed no obvious difference. After incubation, the samples were gently rinsed with copious PBS buffer (>5 mL) to remove the planktonic or loosely attached cells. The rinsed samples were kept in PBS buffer or growth medium at 4 °C for further analysis.

Optical Imaging of Immobilized Cells. Optical imaging was done using either an Olympus BX61 or a Leica TCS SP2 microscope. All of the images were recorded in reflection mode (Figures 2, 4, and 5A,B) or in fluorescence mode (Figures 5C,D and 6). For reflection mode imaging, the samples were imaged using water immersion objective lenses in PBS buffer or 1/2 LB growth medium without staining the bacterial cells. The bright background color of the patterned areas in Figures 2, 4, and 5A,B is due to the Ga^+ focused ion beam etching of the substrate surfaces. For fluorescence mode imaging, the cells were stained using Live/Dead *Bac*Light according to the protocol suggested by the dye manufacturer, and imaged in a PBS buffer using water immersion objective lenses.

X-ray Photoelectron Spectroscopy (XPS). XPS analysis was conducted using a Physical Electronics 5600ci system equipped with monochromatized Al Kα X-rays. Data acquisition and data analysis were performed using RBD AugerScan2 software.

Atomic Force Microscopy (AFM). Images were obtained in tapping mode in air using a Multimode V atomic force microscope from Veeco (Santa Barbara, CA) with an E- or J-type scanner and using NSC18 AFM probes from MikroMasch (Wilsonville, OR).

Field Emission Scanning Electron Microscopy (FESEM). Immobilized and planktonic bacteria deposited on a clean silicon wafer were imaged using a Zeiss SUPRA 55VP system (Carl Zeiss, Germany).

Results and Discussion

Figure 1 shows the method for immobilizing and preparing microscale patterns of live bacterial cells. Depending on the choice of substrate, a clean silicon wafer or a gold substrate surface is first passivated using a PEG layer to prevent nonspecific

(15) Pascual, D. W.; Trunkle, T.; Sura, J. *Infect. Immun.* **2002**, *70* (8), 4273−4281.
(16) Bullifent, H. L.; Griffin, K. F.; Jones, S. M.; Yates, A.; Harrington, L.; Titball, R. W. *Vaccine* **2000**, *18* (24), 2668−2676.
(17) Yang, X. Y.; Hinnebusch, B. J.; Trunkle, T.; Bosio, C. M.; Suo, Z. Y.; Tighe, M.; Harmsen, A.; Becker, T.; Crist, K.; Walters, N.; Avci, R.; Pascual, D. W. *J. Immunol.* **2007**, *178* (2), 1059−1067.
(18) Bansal, T.; Englert, D.; Lee, J.; Hegde, M.; Wood, T. K.; Jayaraman, A. *Infect. Immun.* **2007**, *75* (9), 4597−4607.
(19) Lahiri, J.; Isaacs, L.; Tien, J.; Whitesides, G. M. *Anal. Chem.* **1999**, *71* (4), 777−790.
(20) Jo, S.; Park, K. *Biomaterials* **2000**, *21* (6), 605−616.
(21) Avci, R.; Hagenston, A. M.; Equall, N. L.; Groenewold, G. S.; Gresham, G. L.; Dahl, D. A. *Surf. Interface Anal.* **1999**, *27* (8), 789−796.

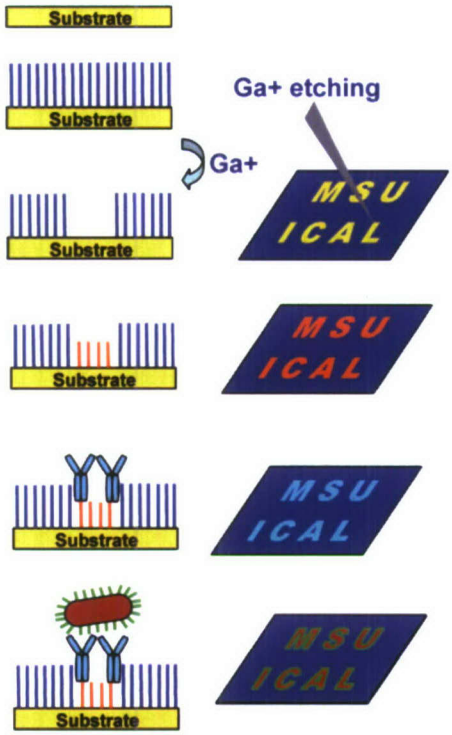


Figure 1. Micropatterning of live bacterial cells. The substrate was first modified with chemicals that inhibit the nonspecific adsorption of proteins (blue bars) and then etched using a programmable focused Ga^+ ion beam. The freshly etched surface was then modified with a cross-linker (orange bars) to link the antibody (cyan Y shapes), raised against the bacterial surface antigens, to the substrate. When the patterned substrate was incubated with the bacterial suspension, the bacterial cells adhered only to the antibody-modified area and thus formed a monolayer of bacterial cell patterns.

adsorption of the antibody.[19,22] The substrate is then patterned using a focused Ga^+ ion beam to remove the preselected portions of the passive layer from the substrate. This is followed by attaching one end of a cross-linker molecule to the sputtered area of the substrate and the other end to an antibody molecule, all via covalent bonds. For silicon, trialkoxysilanes are used to couple the cross-linkers and PEG moieties to the substrate, and for gold substrates thiol-based compounds are used. After rinsing with PBS buffer to remove the excess antibody from the medium, the antibody-patterned substrate is incubated with a bacterial suspension at ambient conditions for a period of about 3 h. The bacterial cells attach only to the antibody-modified areas, so that a micropattern of bacterial cells is achieved.

Optical images of bacterial patterns on gold and silicon substrates obtained while the bacteria were alive are shown in Figure 2. A sharp contrast between the patterned area and the PEG-passivated area is clearly demonstrated in these images. Only sparsely attached cells are observed in the PEG-passivated area, and all of the patterned area is occupied by a dense monolayer of *S.* Typhimurium cells. To obtain such high-quality micropatterns of bacterial cells, attention must be paid to surface preparation and substrate cleanness during the passivation step (Figure 3). Flaws in surface passivation can cause failures of patterning. The failure of passivation facilitates antibody adsorption everywhere on the surface, leading to indiscriminate immobilization (Figure S-1 in the Supporting Information).

The attachment of antibody-immobilized cells was robust enough to resist washing with copious buffer solution for all strains tested. In some experiments, the antibody molecules were allowed to adsorb noncovalently onto the patterned (unpassivated) areas without the cross-linker. Bacterial patterns generated in this way were still resistant to the wash with PBS buffer (though not as strong as covalently coupled antibodies), indicating that the attachment, while not as strong, was robust enough for bacterial immobilization. This is important in practice and avoids the extra step of coupling the antibodies covalently to the substrates. Figure 2A is an example of such immobilization.

The immobilized cells remain viable for at least 6 h in PBS buffer when stored at 4 °C, as evidenced by the fluorescence of the immobilized cells after being stained with viability dyes (Figure 5C). The viability of the immobilized cells in PBS buffer is also supported by their free motion around the locations of immobilization, as shown in Movie S-1 in the Supporting Information. In growth medium, the immobilized cells can last for at least 3 days while retaining their physiological activities, and a longer lifetime can reasonably be expected for the immobilized cells, provided nutrients are available and the antibody molecules remain active. Monitoring the immobilized bacteria under optical microscope showed clearly that the cells were alive and capable of division in the growth medium at ambient conditions (Figure 4 and Movie S-2 in the Supporting Information). It is clear that the bacteria were dividing without hindrance even though they were tied to the substrate through antibody−antigen interactions.

As seen in Figure 2D, cellular resolution of bacterial patterning was achieved in our experiments. Some of the line thicknesses, ~1 μm, in the patterns are comparable to the dimensions of the bacteria, and the bacteria are concentrated along these narrow lines while very few cells are observed outside the lines. In fact, those few cells immobilized outside the patterned areas are also most likely immobilized through antibody−antigen interactions. Passivation is not perfect: some antibodies find their way to the substrate on the defects of the passive layer. These antibodies outside the patterned areas become centers of attachment for the motile bacteria that happen to be at these locations.

Poly-L-lysine has been successfully used to immobilize *E. coli* K-12 strains on abiotic substrates[2,12] as well as to attach eukaryotic cells to glass slides.[23,24] However, it failed to immobilize all the bacterial species we tested, including *E. coli*, *S.* Typhimurium, and *Haemophilus influenzae*. Other approaches to immobilizing bacterial cells, such as an amino-terminated surface,[25] a positively charged polymer,[26] a gelatin-coated surface,[27] and a direct linking of bacterial cells via covalent bonds,[28] proved to be unreliable, inefficient, and irreproducible for the strains we tested. The difficulties associated with these techniques can partially be explained by the fact that bacterial cells have a very small contact area with the substrate surface due to their smaller size relative to eukaryotic cells, typically by 1−2 magnitude orders. Furthermore, a forest of appendages, such as pili and flagella, protruding out of the surface, as shown in the inset of Figure 2A, prevent bacteria from contacting the surface. Additionally, many bacterial species, including *S.* Typhimurium, have a layer of extracellular polymeric substances

(22) Kingshott, P.; Griesser, H. J. *Curr. Opin. Solid State Mater. Sci.* **1999**, *4* (4), 403−412.

(23) Jacobson, B. S.; Branton, D. *Science* **1977**, *195* (4275), 302−304.

(24) Leifer, D.; Lipton, S. A.; Barnstable, C. J.; Masland, R. H. *Science* **1984**, *224* (4646), 303−306.

(25) Arnoldi, M.; Kacher, C. M.; Bauerlein, E.; Radmacher, M.; Fritz, M. *Appl. Phys. A* **1998**, *66*, S613−S617.

(26) Postollec, F.; Norde, W.; de Vries, J.; Busscher, H. J.; van der Mei, H. C. *J. Dent. Res.* **2006**, *85* (3), 231−234.

(27) Doktycz, M. J.; Sullivan, C. J.; Hoyt, P. R.; Pelletier, D. A.; Wu, S.; Allison, D. P. *Ultramicroscopy* **2003**, *97* (1−4), 209−216.

(28) Camesano, T. A.; Natan, M. J.; Logan, B. E. *Langmuir* **2000**, *16* (10), 4563−4572.

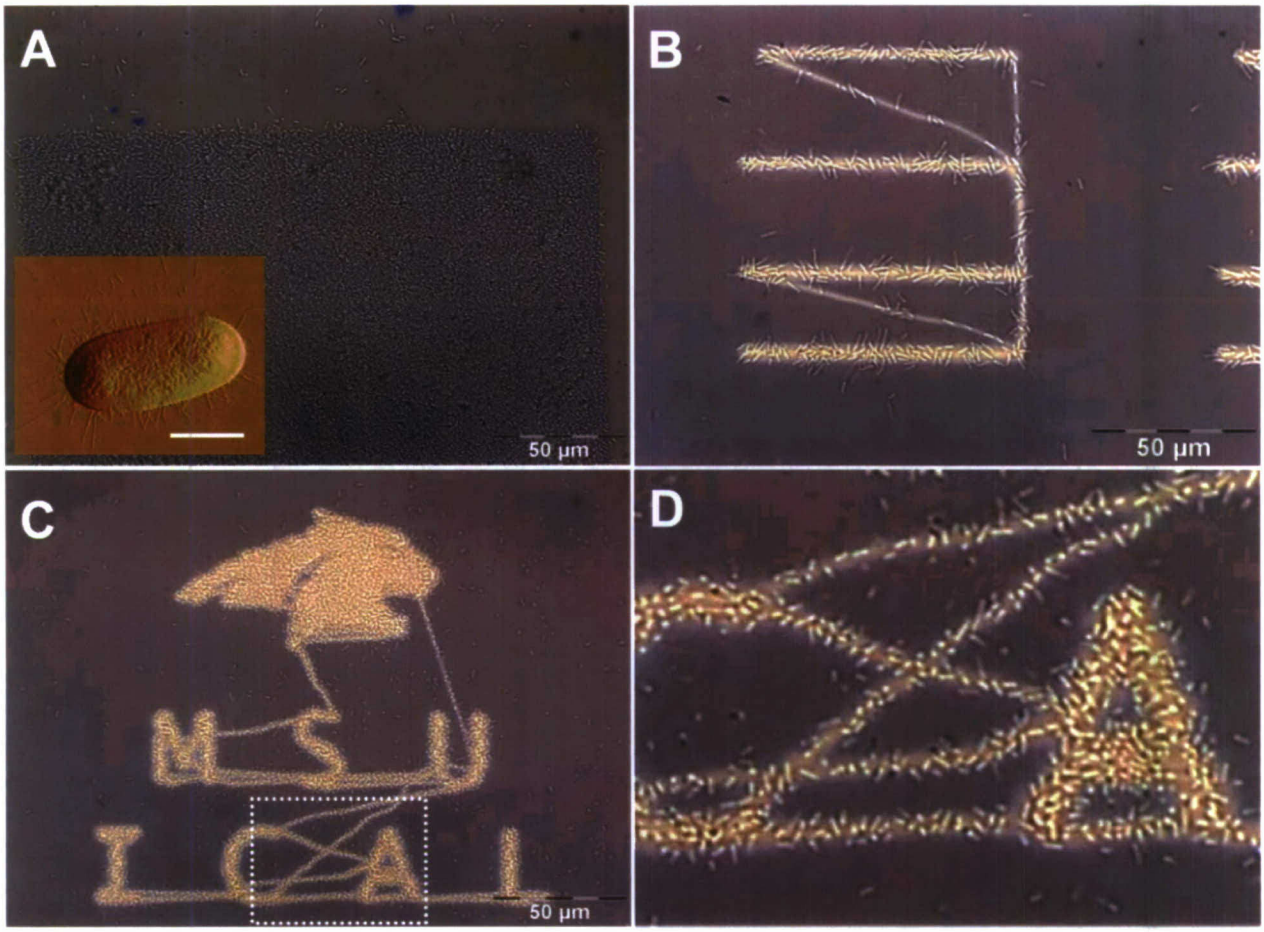


Figure 2. Micropatterns of live *S.* Typhimurium cells immobilized on substrates etched by a focused ion beam. (A) Square pattern on gold. The inset in (A) shows a high-resolution atomic force microscope image of a typical *S.* Typhimurium Δ*asd*::*kan*^{R} H71-pHC, obtained in air, revealing the CFA/I fimbriae (scale bar: 1 μm). (B) Line pattern on silicon. (C) Bobcat mascot and MSU-ICAL logo on silicon. (D) Enlargement of the area within the white dotted box in (C), demonstrating the cellular resolution of cell patterning.

surrounding the bacteria (capsular EPS)[29] which prevents the bacteria from direct contact with the surface.

The choice of antibody is critical to the success of bacterial immobilization. The purity of the antibody is crucial; an affinity-purified antibody is preferred, because the serum contains a considerable amount of proteins other than the desired antibody, for example, serum albumin, which will compete for the bioactive cross-linker sites on the surface during the covalent linking of antibody proteins to the substrate. This will then mostly coat the surface with serum albumin, reducing the fraction of the surface covered with antibody molecules. Therefore, the use of a serum without affinity purification will yield a low density of antibody molecules on the substrate and thus give rise to a low density of immobilized bacteria. This assumption is supported by our early attempts to immobilize bacterial cells through unpurified anti-CFA/I serum, which only afforded unreliable poor immobilization. The efficiency of bacterial immobilization also depends on the choice of antibody−antigen pair. Bacterial species express a large number of surface antigens which play various roles in bacterial virulence and adhesion.[30] Antibodies against many of these surface antigens are commercially available, but we recommend that those appendages that protrude outside the cell walls and do not have rapid movements independent of the bacteria be considered. For example, anti-flagellin is not a good choice for leashing bacteria to substrates because flagella, as the bacterial motility organ, rotate with very high speeds, exceeding 10 000 rpm at 35 °C,[31] vastly reducing the chance of antibody−antigen interactions. The interaction between an antibody−antigen pair is fairly weak: a force of ~50 pN is required to break the antibody−antigen interaction,[32] which is more than 1 order of magnitude smaller than a covalent bond (>1 nN). At this time, we do not know how much force a flagellum imparts to a *S.* Typhimurium cell, but, to immobilize a bacterium, we hypothesize at least one antibody−antigen interaction is required to hold the bacterium in place. For the reasons discussed above, in our experiments, the CFA/I fimbriae were chosen as the target antigen because (1) they are long and protrude outside the cell body (inset of Figure 2A), (2) the fimbriae are expressed in *S.* Typhimurium strains in abundance, and (3) these appendages do not rotate rapidly independently of the bacteria as do the flagella. A typical *S.* Typhimurium cell will fit into a 2 × 1 μm^2 area (inset Figure 2A), which predicts a maximum packing density of about one bacterium per 2 μm^2. This value is in good agreement with our observations conducted under optical microscope (Figure 2A,C and the Supporting Information).

The unique features of the bacterial cells patterned in this way are the stability and self-sustaining capability of the patterns. As shown in Figure 4 and Movie S-2 in the Supporting Information,

(29) Suo, Z. Y.; Yang, X. Y.; Avci, R.; Kellerman, L.; Pascual, D. W.; Fries, M.; Steele, A. *Langmuir* **2007**, *23* (3), 1365−1374.

(30) Ofek, I.; Hasty, D. L.; Doyle, R. J. *Bacterial adhesion to animal cells and tissues*; ASM Press: Washington, D.C., 2003.

(31) Magariyama, Y.; Sugiyama, S.; Kudo, S. *FEMS Microbiol. Lett.* **2001**, *199* (1), 125−129.

(32) Allison, D. P.; Hinterdorfer, P.; Han, W. H. *Curr. Opin. Biotechnol.* **2002**, *13* (1), 47−51.

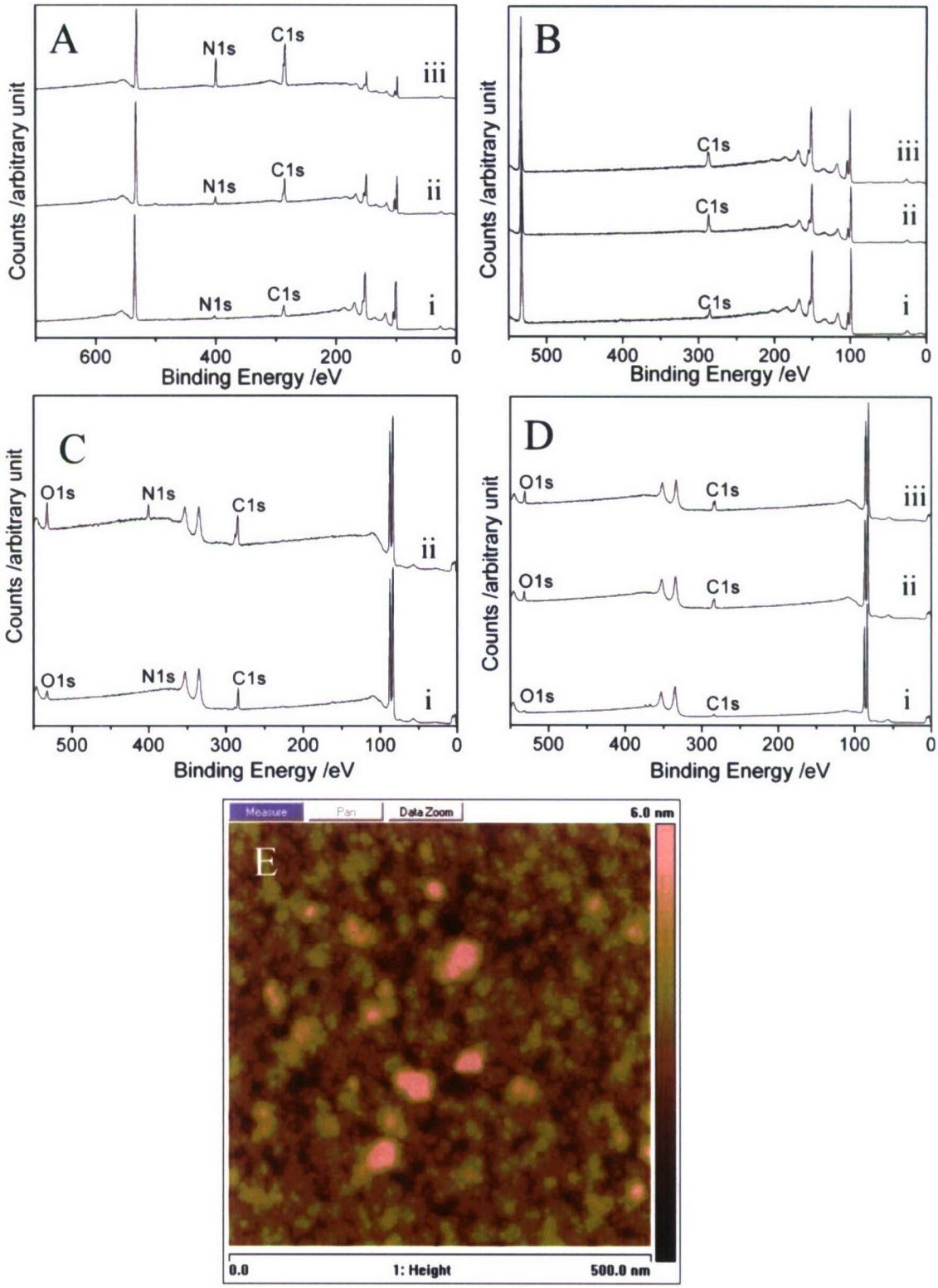


Figure 3. Characterization of silicon and gold substrates. (A) XPS spectra of silicon chips (i) modified with APTES; (ii) modified with APTES and BMPS; and (iii) modified with antibody. No N1s signal was observed for unmodified silicon, and both C1s and N1s signals increased as APTES, BMPS, and antibody were linked to the silicon surface (from i to iii), implying that the antibody was successfully bonded to the silicon surface with APTES/BMPS as the cross-linkers. (B) XPS spectra of silicon chips (i) cleaned by O_3 plasma; (ii) passivated using PEG-silane; and (iii) passivated and incubated with anti-CFA/I antibody. The increase of the C1s peak from (i) to (ii) suggests the success of the passivation of the silicon surface using PEG-silane, while the absence of the N1s peak in (iii) confirms the resistance against antibody adsorption of the passivated silicon surface. (C) XPS spectra of gold substrates (i) modified with 11-MUDA and (ii) with anti-CFA/I antibody cross-linked. For the gold surface, C1s, N1s, and O1s signals were used to monitor the substrate modification. The prominent N1s signal in (ii) and the increase of O1s from (i) to (ii) indicate the attachment of the antibody. (D) XPS spectra of gold substrates (i) cleaned by O_3 plasma; (ii) passivated with SPT-11; and (iii) passivated and incubated with anti-CFA/I antibody. The increases in the C1s and O1s peaks from (i) to (ii) suggest the success of the passivation of the gold surface using SPT-11, while the absence of the N1s peak in (iii) confirms the resistance against antibody adsorption of the passivated gold surface. (E) AFM image of antibody molecules covalently linked to the silicon surface. Antibody molecules linked to the gold surface were not imaged by AFM because of the interference from the gold nanograins.

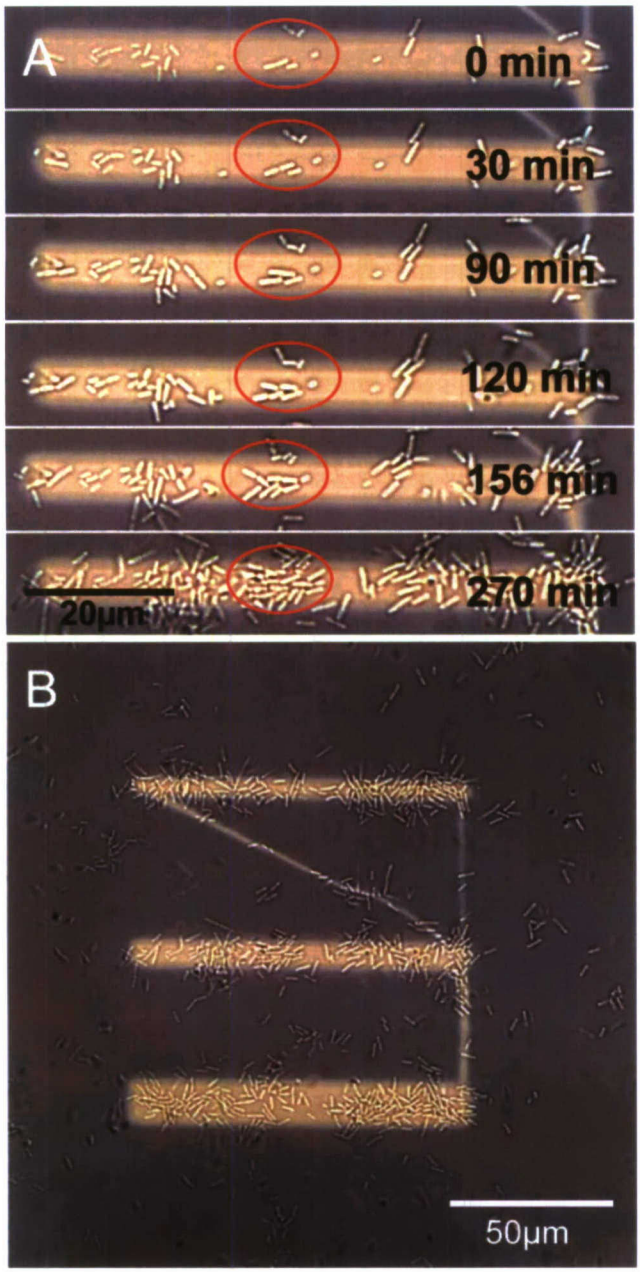


Figure 4. (A) Selected time lapse images over an incubation period of 270 min, showing the regeneration of immobilized *S.* Typhimurium on a patterned silicon substrate that was incubated in growth medium at room temperature. The red ovals highlight the dividing cells. (B) Time lapse image at 270 min showing that newly divided cells only occupy the available patterned areas and that excess cells are released into the growth medium and can be removed by gently rinsing.

the immobilized cells are capable of dividing and regenerating. The daughter cells either stay attached to the antibody-modified area or enter the growth medium after the antibody-modified area is fully occupied. Therefore, the patterns are always covered by a monolayer of live cells and do not change physical dimensions as a result of cell division. In a flow reactor, cells released into the liquid phase will be carried away and the self-refreshing bacterial pattern can last as long as the antibody remains active, provided that the medium has the nutrients to sustain life. This is particularly useful for those applications, such as water quality control, that require continuous monitoring of pollutants as an environmental sensor. These features are not available for patterns prepared by mechanically transferred cells on an agar plate;[4,5] for sensors fabricated by incorporating cells in alginate,[33] agar,[34] or sol−gel matrices;[35] or for arrays constructed by physical entrapment into microwells.[8] For these patterns, bacteria are held on the substrates or inside the microwells with weak forces, and the cells do not stay fixed to the patterned areas for long periods of time. These patterns will eventually degrade and disintegrate when exposed to a liquid medium or a flow reactor.

The observation of patterned bacteria on antibody-modified substrates allows probing of the individual or collective behavior of the bacteria. For example, we can follow the increase in surface population density of bacteria incubated in a growth medium. Such observation in our experiments yielded a surprising result, as shown in Figure 5A,B, in which bacteria immobilized in a lying-down orientation took a standing-up orientation as their density increased. This standing-up orientation of the crowded cells was also confirmed by laser scanning confocal microscopy (LSCM) images (Figure 5 C,D). Although this behavior is not well understood at this time, it might be related to the depletion of nutrients at the crowded bacterial positions and to the struggle of the bacteria to move away from their immobilized positions. This hypothesis is supported by the observation of an excess number of flagella produced by the immobilized bacteria (Supporting Information Figure S-2), presumably in an effort to free themselves from their positions.

The use of purified antibodies targets specific antibody−antigen interactions; hence, only bacteria with the targeted antigens are immobilized. This specificity of immobilization can be used in many applications, including sorting organisms from a mixed culture. A successful sorting of *S.* Typhimurium (green) from a culture mixed with *E. coli* (red) is demonstrated in Figure 6. The mixture (Figure 6A) was allowed to interact at ambient conditions for only 2 h with a substrate patterned in a checkerboard geometry with antibodies specific to *S.* Typhimurium. Figure 6B shows only the desired organism (*S.* Typhimurium) immobilized on the antibody patterns: *E. coli* cells were washed off the substrate. Sorting a specific bacterial strain typically takes weeks to months, because the mixed culture is subjected to the repeated inoculation and growth of sequentially diluted cultures on preselected agar plates.[36]

Could bacteria be linked directly to a substrate without a need for an antibody? The advantage of this method, if successful, would be to avoid the necessity of finding an antibody for a given bacterial strain. The disadvantage would be that all the proteins, and possibly other biomolecules, in the medium would have a chance at immobilization; hence, the bacterial specificity would be lost. We tested this idea by linking bacterial cells directly to a substrate through covalent coupling between bacterial membrane proteins and cross-linkers on the substrate, as was done with the antibodies. The activated substrates, with either carboxyl groups or maleimido groups, were directly incubated with a suspension of bacterial cells in the growth medium and also in the PBS buffer solution for ~15 h under ambient conditions. The results (not shown) indicate only sparsely attached cells populating the surface. We hypothesize the reason for the failure of direct immobilization is that the activated surface was blocked by proteins in the bacterial suspension before the bacterial cells could reach the surface. These proteins are components of

(33) Polyak, B.; Bassis, E.; Novodvorets, A.; Belkin, S.; Marks, R. S. *Sens. Actuators, B* **2001**, *74* (1−3), 18−26.
(34) Mbeunkui, F.; Richaud, C.; Etienne, A. L.; Schmid, R. D.; Bachmann, T. T. *Appl. Microbiol. Biotechnol.* **2002**, *60* (3), 306−312.
(35) Premkumar, J. R.; Rosen, R.; Belkin, S.; Lev, O. *Anal. Chim. Acta* **2002**, *462* (1), 11−23.
(36) Eaton, A. D.; Clesceri, L. S.; Rice, E. W.; Greenberg, A. E. *Standard methods for the examination of water and wastewater*, 21st ed.; American Public Health Association: Washington, D.C., 2005.

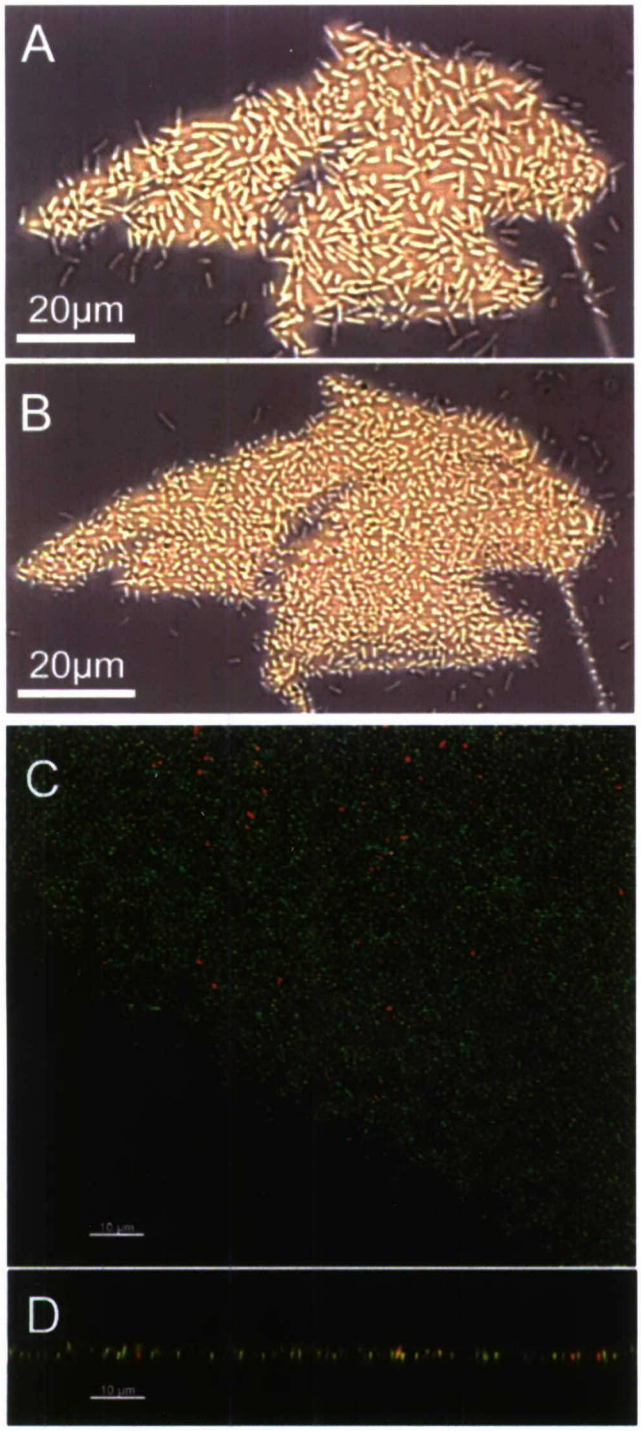


Figure 5. Immobilized *S.* Typhimurium cells change their orientations according to the surface population density on the substrate. Images (A) and (B) correspond to a sample incubated in growth medium at 37 °C for 3 and 15 h, respectively. Notice that while the cells are lying down in part (A), they appear to stand up in (B), when the surface is crowded (see Figure 5D). LSCM image of a sample incubated in growth medium for 15 h and then stored in PBS buffer at 4 °C for 6 h and stained with viability stains (C) indicating that the majority of the cells were alive. The reconstituted *Z*-section image of the immobilized cells (D) further confirms that a majority of the cells took a standing-up orientation.

the growth medium or are secreted by the bacteria in the suspension. It is difficult to remove these proteins completely, even by washing the cells with PBS buffer repeatedly. Furthermore, the capsular EPS layer surrounding each bacterium would prevent the bacterial surface proteins from making direct contact with the activated substrate surface.

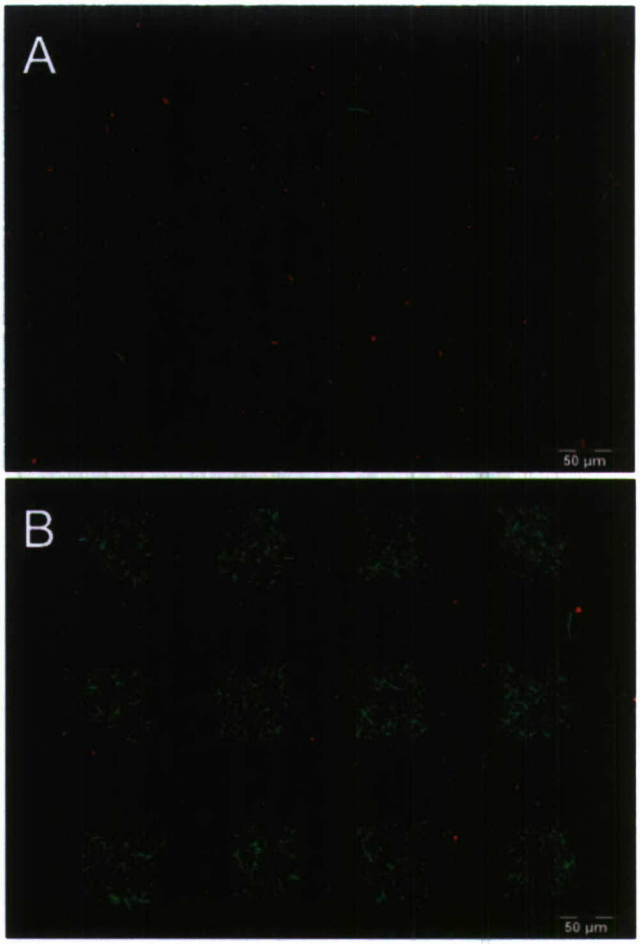


Figure 6. Sorting *S.* Typhimurium cells from a mixture of *S.* Typhimurium and *E. coli*. (A) Epifluorescence image of a mixed culture of *S.* Typhimurium expressing GFP and *E. coli* expressing RFP. (B) Epifluorescence image of the sorted cells on silicon using a checkerboard microarray pattern of an antibody highly specific to the CFA/I fimbriae of *S.* Typhimurium.

Summary

Our results demonstrate that cellular resolution can be achieved for the preparation of self-renewing bacterial patterns on abiotic substrates through antibody−antigen binding with high efficiency and controllability. The efficiency of the immobilization is attributed to the careful selection and assembly of the substrate surface chemistry at each step and to the affinity-purified antibody raised specifically against a carefully targeted antigen protruding outside the cell surface of the bacteria, the CFA/I fimbriae of *S.* Typhimurium. Patterned cells remain viable under this *immobilized* condition and are capable of reproducing. The technique is readily applicable to sorting specific bacteria from a mixed culture within hours as compared to the standard bacteria purification methods, which typically take days.

Acknowledgment. This work is funded by NASA-EPSCOR under Grant NCC5-579, by MSU CBIN, by U.S. Public Service Grant AI-41123, and by Montana Agricultural Station and USDA Formula funds. We thank Mr. M. Deliorman for his help on TOFSIMS and Dr. B. Pitts for her help on optical microscope imaging.

Supporting Information Available: Images of immobilized *S.* Typhimurium cells and the flagella expressed by the immobilized cells (Figures S-1 and S-2), and movie clips (Movies S-1 and S-2) showing the motion and dividing of immobilized cells. This material is available free of charge via the Internet at http://pubs.acs.org.

LA7038653

doi:10.1016/j.jmb.2007.10.044 J. Mol. Biol. (2008) 375, 581–594

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Conformational Equilibria and Rates of Localized Motion within Hepatitis B Virus Capsids

Jonathan K. Hilmer, Adam Zlotnick and Brian Bothner*

Montana State University, Department of Chemistry and Biochemistry, Chemistry and Biochemistry Building, Bozeman, MT 59717, USA

Received 19 May 2007; received in revised form 8 October 2007; accepted 9 October 2007
Available online
22 October 2007

Functional analysis of hepatitis B virus (HBV) core particles has associated a number of biological roles with the C terminus of the capsid protein. One set of functions require the C terminus to be on the exterior of the capsid, while others place this domain on the interior. According to the crystal structure of the capsid, this segment is strictly internal to the capsid shell and buried at a protein–protein interface. Using kinetic hydrolysis, a form of protease digestion assayed by SDS–PAGE and mass spectrometry, the structurally and biologically important C-terminal region of HBV capsid protein assembly domain (Cp149, residues 1–149) has been shown to be dynamic in both dimer and capsid forms. HBV is an enveloped virus with a $T=4$ icosahedral core that is composed of 120 copies of a homodimer capsid protein. Free dimer and assembled capsid forms of the protein are readily hydrolyzed by trypsin and thermolysin, around residues 127–128, indicating that this region is dynamic and exposed to the capsid surface. The measured conformational equilibria have an opposite temperature dependence between free dimer and assembled capsid. This work helps to explain the previously described allosteric regulation of assembly and functional properties of a buried domain. These observations make a critical connection between structure, dynamics, and function: made possible by the first quantitative measurements of conformational equilibria and rates of conversion between protein conformers for a megaDalton complex.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: protein dynamics; Hepatitis B virus; proteolysis; supramolecular complexes; virus assembly

Edited by M. F. Summers

Introduction

In solution, many proteins exist as an ensemble of conformations. The population distribution of the ensemble is dependent on the barriers between energy minima on the multi-dimensional energy landscape of the protein. In recent years, the strong connection between dynamics and function has made the study of protein dynamics a critical research frontier.[1–3] The requirement of local conformational plasticity in biological function is well documented for large complexes such as the ribosome,[4] nuclear pores,[5] and virus particles.[6,7] A range of biophysical approaches can be used to describe the energy landscape of a protein (i.e. the distributions of different conformations and the rates of exchange), and this has been done for a number of monomeric proteins.[8–10] However, due to technical challenges, detailed mapping is not available for the dynamics of viruses, or any other, megaDalton complexes. Enzymatic hydrolysis has a long history as an effective tool for distinguishing between folded and unfolded proteins and in the identification of dynamic regions. Here we demonstrate that when applied rigorously, kinetic hydrolysis allows the determination of conformational equilibria and rate constants for megaDalton nucleoprotein structures.

More than 350 million people suffer from chronic infection with hepatitis B virus (HBV). It is one of the major contributing factors to hepatocellular carcinoma and cirrhosis. HBV is an enveloped virus with an icosahedral core. The majority of cores are ~350 Å in diameter[11] and are comprised of 120 core protein homodimers arranged with $T=4$ icosahedral symmetry[12] (Figure 1). The core protein

*Corresponding author. E-mail address: bbothner@chemistry.montana.edu.

Abbreviations used: HBV, hepatitis B virus; HDX, hydrogen–deuterium exchange; MALDI, matrix-assisted laser desorption ionization.

0022-2836/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.

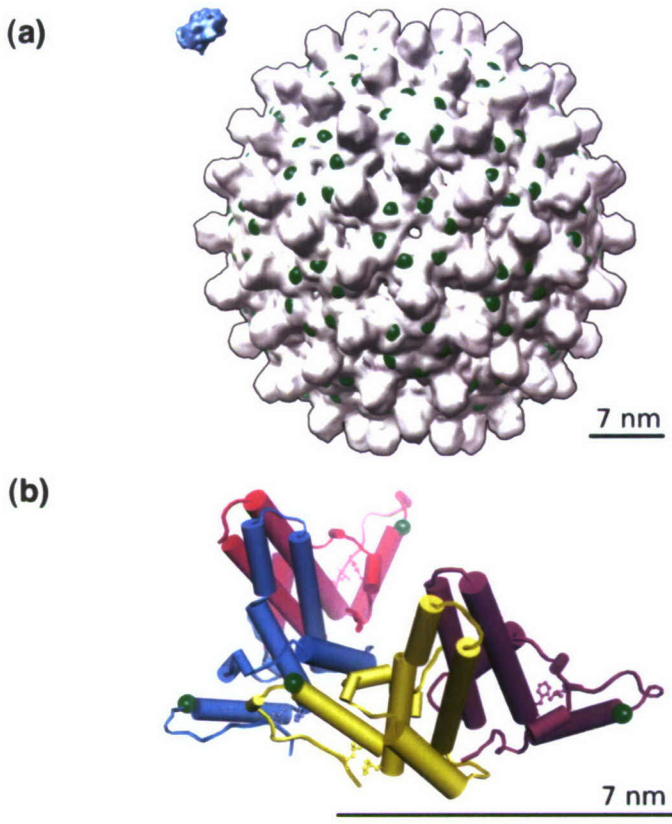


Figure 1. Hepatitis B capsid. (a) The HBV capsid has $T = 4$ symmetry and is assembled from 120 copies of Cp149 dimer. A hexamer of dimers surrounds the quasi-6-fold (icosahedral 2-fold) axis. The C terminus of the protein is located at the 5-fold and 6-fold symmetry axes, and in all structural models to date the C terminus of the protein projects inward. A model of trypsin is included for scale (blue, upper left), and the observed cleavage site is indicated with exaggerated scale (green spheres). A scale bar of 7.0 nm is shown in the lower-right for (a) and (b). (b) The asymmetric unit from the capsid crystal structure. Each dimer contains two copies of Cp149, which meet together with a 30 Å tall helical interface. Contact between dimers occurs in a limited region spatially centered around the end of helix α-5. The cleavage site at residues R127/T128 is indicated by small green spheres and residue L140 on the C-terminal coil is shown where it meets F110 and Y118 as part of the hydrophobic interactions stabilizing the folded conformation observed in the crystal structure. (c) The amino acid sequence for Cp149. The cleavage site is shown in green, and a putative ubiquitin ligase binding site is indicated in blue.

(HBcAg) has 183 amino acids and can be divided into an assembly domain, residues 1–149 (Figure 1(c)), and a C-terminal 34 amino acid RNA-binding domain. Only the assembly domain of core protein is required to form the capsid. HBV also produces a variant of the protein (HBeAg) that includes an additional ten amino acids at the N terminus but terminates at position 149. The 149 residue construct from residues 1 to 149 (Cp149), lacking the C-terminal domain, forms well-ordered structures[13,14] that are indistinguishable from the cores of native virions.[15] The assembly domain dimer ($Cp149_2$) is the basic unit of capsid assembly. It forms spontaneously at translation[16] and requires a denaturing concentration of chaotrope to dissociate.[17] The structure of the dimer is dominated by the interface between monomers, a four-helix bundle which projects 30 Å from the plane of the capsid surface (Figure 1(a)), and is comprised of a helix-turn-helix from each monomer (Figure 1(b)).[13,14,18,19] It should be noted that all structural models of $Cp149_2$ are based on intact particles. $Cp149_2$ can be expressed as a recombinant protein in *Escherichia coli*, where it spontaneously forms capsids: these capsids can be dissociated into dimers and reassembled.[20] The assembly process of recombinant particles has been well-characterized from both kinetic and thermodynamic standpoints, making HBV one of the best understood model systems for studying virus capsid assembly.[17,21–24] Capsid assembly is a function of pH, temperature, and ionic strength.[23,24] Assembly models predict that the dimer undergoes a transition between inert and assembly-competent conformations.[23] A number of small molecules act as effectors that enhance or inhibit the assembly process.[21,25–31] The best understood of these, HAP1, enhances assembly and particle stability, at least in part, by stabilizing the conformation of the C terminus of the assembly domain.[14]

Capsid dynamics are likely critical to HBV function. For example, the C-terminal domain has interactions outside the capsid, playing a major role in intracellular trafficking,[32] even though in structural models it is within the capsid interior.[13,18,19] Phosphorylation of the C-terminal domain is required for nuclear localization,[33–36] while loss of phosphorylation correlates with particle secretion.[37] It has also been suggested that a PPAY sequence near the C terminus of the assembly domain is critical for trafficking to the endoplasmic reticulum (ER).[38] These results all suggest that, at least transiently, the C-terminal domain of core protein is exposed on the exterior of the HBV particle. However, the C-terminal domain is also associated with interactions that take place on the capsid

interior. Phosphorylation of the C-terminal domain is required for specific packaging of viral RNA,[39–41] and the pattern of phosphorylation directly effects reverse transcription, which takes place within the assembled capsid.[34,41,42] Together, these data create a paradox where the C-terminal domain is localized to the capsid interior but is functionally active on the exterior. This can be resolved if the C terminus is dynamic; however, quantitative evidence has thus far been lacking.

Kinetic hydrolysis is well suited to the study of stability and dynamics in the context of large complexes such as the HBV particle. In this technique, a protein sample is hydrolyzed using a proteolytic enzyme under carefully controlled conditions. Based upon the well-established theory that proteolysis only occurs in exposed regions lacking well-defined structure,[43–46] this technique has long been used to isolate individual domains for structural analysis,[47] identify dynamic protein regions,[48,49] and more recently as a tool to quantitatively probe the high energy states of proteins.[50]

The fundamental kinetic model involves dynamic motion between a protected "native" state and a partially unfolded, "open" conformation, which is cleaved by protease at a known rate:

$$Closed \underset{k_{close}}{\overset{k_{open}}{\rightleftharpoons}} Open \xrightarrow{k_{cleave}} Cleaved \qquad (1)$$

By varying enzyme concentration and quantifying its effect on the overall reaction rate, the equilibrium and rate of transition between closed and open forms of the protein can be assayed. By definition, the protein in the open conformation behaves as an intrinsically disordered polypeptide and is cleaved by the protease following standard Michaelis–Menten kinetics.[50] This gives rise to a simple kinetic model where disappearance of intact protein depends on the concentration of protein in the open conformation, catalytic constants (k_{cat}/K_M) and enzyme concentration (equation (8)). Although systematically similar to hydrogen–deuterium exchange (HDX),[51,52] the scale of motion required for detection by proteolytic hydrolysis of folded proteins is much greater. The exact deformation from the protected "native" state will vary depending on local constraints, but a typical motion will involve a substantial displacement and unfolding of any structure within two to four amino acids on each side of the cleavage location.[44] This large-scale motion is expected to be rare compared to the smaller transitions required for amide proton exchange in HDX. Protein dynamics on the nanometer scale are known to be involved in complex formation,[53] viral capsid breathing[49,54–56] and viral docking and entry events.[57]

Here we present the first quantitative measurements of conformational equilibria and rates of conversion between conformers for a megaDalton complex, the HBV capsid. Our results indicate that the C terminus of Cp149 is dynamic and transiently exposed on the surface of the capsid, in contrast to its position in crystal and cryo-electron microscopy structures. The exposure of this region is highly dependent on the assembly state of the protein as well as the temperature. The enthalpy and entropy of the opening transition for the C terminus reverse sign as a function of assembly, identifying this region as a thermodynamic molecular switch. These results suggest that regulation of assembly may involve the dynamic motion of the C-terminal region of $Cp149_2$.

Results

Identification of dynamic domains

HBV Cp149 has two assembly states: dimer and capsid. Once assembled into a capsid, there is a substantial hysteresis to disassembly, making it possible to conduct separate experiments on dimer and capsid forms under identical solution conditions.[17] For proteolytic reactions, capsid and dimer were diluted into reaction buffer conditions immediately prior to the initiation of the reaction, and the reactions never exceeded 4 h in length. Using nearly identical dilution-based procedures to study disassembly reactions, capsids at similar concentrations remain stable for at least five days.[17] Size-exclusion chromatography and dynamic light scattering were used to confirm that capsids remained intact throughout the initial phase of digestion, which was defined as the time period with at least 70% remaining uncleaved protein. Capsids were found to be fully stable at this level of cleavage, with no detectable dissociation (Figure 2(b)). All of the kinetic hydrolysis experiments conducted on assembled capsid were carefully designed to only include data collected before the appearance of degraded particles.

To probe for dynamic regions in the core protein and to investigate the effect of particle assembly on stability and dynamics, we first tested different proteases and analyzed samples from time-course reactions by SDS–PAGE. Both trypsin and thermolysin degrade the intact protein in a measurable progression (Figure 3(a) and (b), black triangles). In the initial reaction phase, both proteases produced a single primary product band approximately 3 kDa smaller than the intact protein (Figure 3(a) and (b) gray triangles). A second product band arose after further digestion that was another ~2 kDa smaller as judged by SDS–PAGE: this second band has weak intensity (never greater than 10% total composition) and has greatly delayed appearance, especially for assembled capsid. These results indicated that there was a dynamic or unstructured region in the protein and that the initial reaction product was relatively stable. In addition, the same site was accessible to protease in dimer and capsid. Other cleavage products were subsequent to the primary cleavage and evident only after more aggressive proteolysis. These late-stage digestion products showed differ-

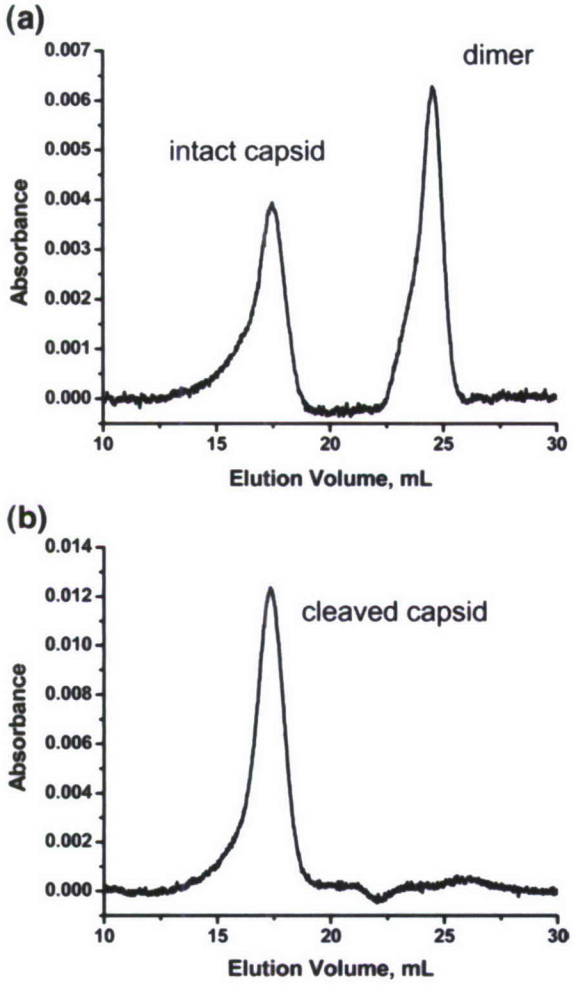


Figure 2. HBV capsid protein maintains integrity during proteolysis. (a) A test mixture of capsid and dimer stock separated with size exclusion chromatography, using the same buffer conditions as the proteolysis reactions. (b) The purified capsid stock was digested with the standard protocol to the level of 70% remaining uncleaved protein, then run over the size exclusion column. No dimer signal is seen, and the capsid intensity is not diminished significantly. The only background subtraction used for (a) and (b) is constant linear correction.

ing patterns between dimer and capsid (data not shown).

Gel densitometry analysis of proteolytic reactions revealed a consistent first-order exponential decay trend for the parent protein (Figure 3(c), circles). Proteolytic cleavage rate constants (k_{exp}) were measured from 19 °C to 37 °C for both HBV capsid and dimer. The observed rate of digestion for dimer is up to 20-fold faster than capsid, necessitating adjustments to sampling times in order to accurately characterize the first 30% of digestion for both species (Figure 3(d), circles). First-order behavior was seen at all temperatures and the change in reaction rate with temperature followed the well-established temperature sensitivity for trypsin.[58,59] In the presence of 5.0 μg/ml of trypsin at 25 °C, the observed rate constant for hydrolysis of dimer is approximately 0.013 min^{-1}, with a half-life of ~53 min.

The precise location of proteolytic cleavage was identified using mass spectrometry. Matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) were employed to analyze both quenched and online reactions. The only significant peptide product mapped to a C-terminal fragment of Cp149 comprised of residues 128–149 (measured: 2333.263(±0.012) Da, predicted: 2333.263 Da). Gels showing the intact protein and primary cleavage product confirm the presence of a species matching the size of amino acid residues 1–127 of Cp149. For quantitative reactions, no evidence for alternative cleavage sites was found by mass spectrometry. The lack of trypsin cleavage sites within the 128–149 peptide rules out enzymatic competition from the released product. Mass spectrometry and gel-based results are in agreement that, during the initial stages of digestion proteolysis proceeds exclusively *via* the R127/T128 cleavage (Figure 1(c)).

Kinetic model

Based upon the observed first-order kinetics of proteolytic cleavage and the single cleavage product, we adopted the two-state model described in equation (1) as a working hypothesis for core protein cleavage (for detailed definitions of equations and constants see Materials and Methods). Because trypsin requires a region of unfolded amino acids bordering the cleavage location,[44] a protein which contains secondary structure or steric protection of a cleavage site will be resistant to proteolysis: this is defined as the "closed" state. Random or concerted motion has the potential to transiently remove this protection in the normal course of a protein's dynamic behavior in solution. Upon local or global unfolding (de-protection) the cleavage site adopts the "open" state and is exposed to hydrolytic attack. This two-state model has had wide acceptance for a variety of protein systems,[50,60–62] and the observed kinetic trends for Cp149 validate the two-state model for this application (for details, see Discussion).

Because trypsin is highly specific in not only the P and P′ sites but also P2–P3′, the pentameric peptide Abz-IRTPP-Tyr(3-NO_2) was used as a model substrate to determine the k_{cleave} constant, the intrinsic cleavage constant for an unstructured peptide at a given enzyme concentration, under the exact conditions used with Cp149. This constant contains a k_{cat}/K_M component that describes the catalytic efficiency on a specific peptide backbone in the open state (equation (6)). By keeping substrate concentrations sufficiently low, k_{cleave} is linearly dependent on enzyme concentration. Fluorometric assays of trypsin activity with the model peptide were accurate at nanomolar concentration, giving k_{cat}/K_M values for all appropriate temperatures (Table 1). The measured K_M value of 20 μM is well above the sub-

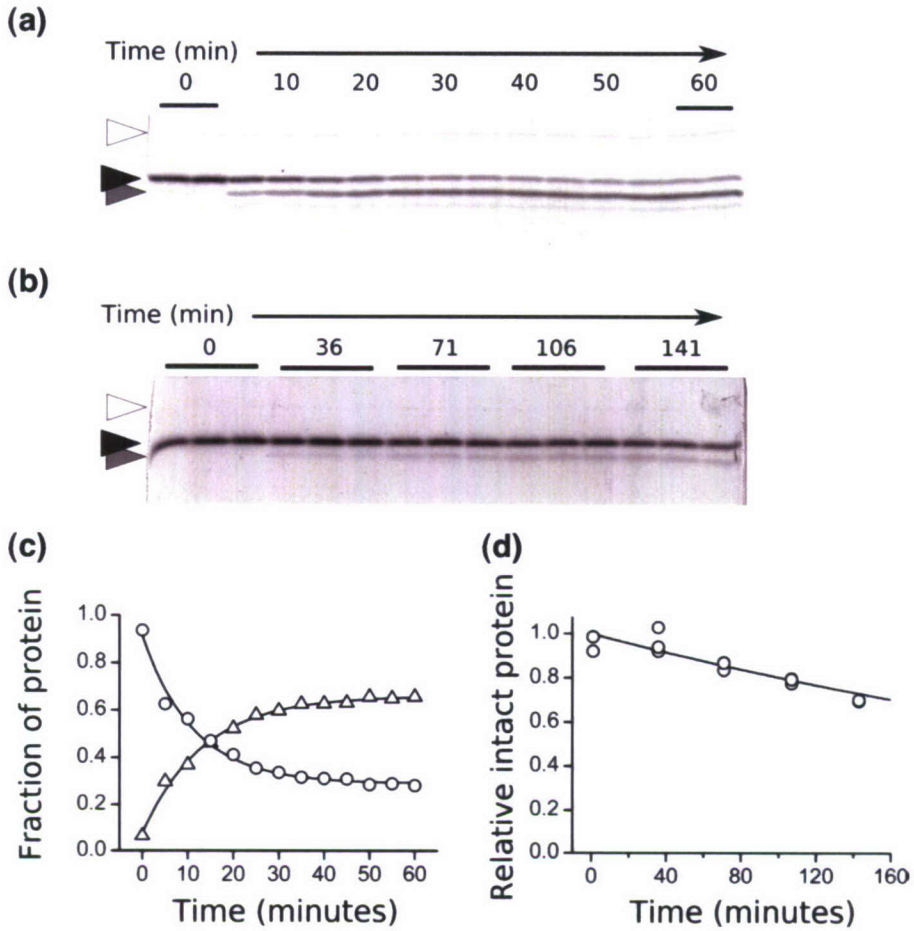


Figure 3. Hydrolysis of HBV capsid protein monitored with SDS–PAGE. (a) The progression of proteolysis in the presence of high concentrations of enzyme (37 °C, dimer). Trypsin (open arrow), intact Cp149 (black arrow), and the cleavage product (gray arrow) are indicated. Alternating time samples (5 min spacing) are labeled, and first and last lanes are run in duplicate. (b) Experiments for determination of observed rate constants (k_{exp}) are restricted to less than 30% cleavage (25 °C, capsid). Multiple sampling replicates for each time point are run to estimate quantitation error. (c) Quantitation of the bands shown in (a). The disappearance of the intact protein (circles) and appearance of product (triangles) follow first-order kinetics. (d) Gel band intensities from (b) were fit to an exponential decay to determine the rate.

strate concentration in the kinetic proteolytic assays (0.3 nM–300 nM), which validates the first-order simplification of Michaelis–Menten kinetics and justifies the treatment of k_{cleave} in the kinetic model as a constant that scales linearly with enzyme. Kinetic characterization of trypsin using the synthetic peptide revealed no significant decrease in k_{cat}/K_M over an 8 h span, a test for any potential complication from enzyme self proteolysis or other degradation during the course of Cp149 cleavage experiments.

For both dimer and capsid forms of Cp149, the maximum k_{exp} rate constant observed with high concentrations of protease (Figure 4(a)) correlated with opening-limited conditions, where the protein is cleaved as soon as it reaches an open state. Under opening-limited conditions, the cleavage rate is equal to the unfolding rate at a given cleavage site

Table 1. Kinetic and thermodynamic parameters for the IRTPP cleavage site

Temperature (°C)	19	25	31	37
Peptide k_{cat}/K_M ($min^{-1}M^{-1}$)	$7.66 \pm 0.23 \times 10^{6}$	$8.92 \pm 0.04 \times 10^{6}$	$10.6 \pm 0.26 \times 10^{6}$	$12.5 \pm 0.46 \times 10^{6}$
Dimer k_{op}, (min^{-1})	$3.53 \pm 0.38 \times 10^{-3}$	$5.08 \pm 0.51 \times 10^{-3}$	$7.60 \pm 0.70 \times 10^{-3}$	$1.19 \pm 0.13 \times 10^{-2}$
Capsid k_{op}, (min^{-1})	$2.44 \pm 0.24 \times 10^{-3}$	$4.57 \pm 0.78 \times 10^{-3}$	$1.13 \pm 0.10 \times 10^{-2}$	$1.29 \pm 0.07 \times 10^{-2}$
Dimer (K_{op})	$1.33 \pm 0.83 \times 10^{-2}$	$7.54 \pm 0.26 \times 10^{-3}$	$6.34 \pm 0.80 \times 10^{-3}$	$2.59 \pm 1.03 \times 10^{-3}$
Capsid (K_{op})	$5.12 \pm 2.28 \times 10^{-3}$	$1.78 \pm 1.56 \times 10^{-3}$	$2.90 \pm 0.62 \times 10^{-3}$	$1.24 \pm 0.30 \times 10^{-2}$
Dimer opening: $\Delta G°$ (kJ mol^{-1})	11.59	12.11	12.79	15.35
Capsid opening: $\Delta G°$ (kJ mol^{-1})	18.40	15.68	14.77	11.31

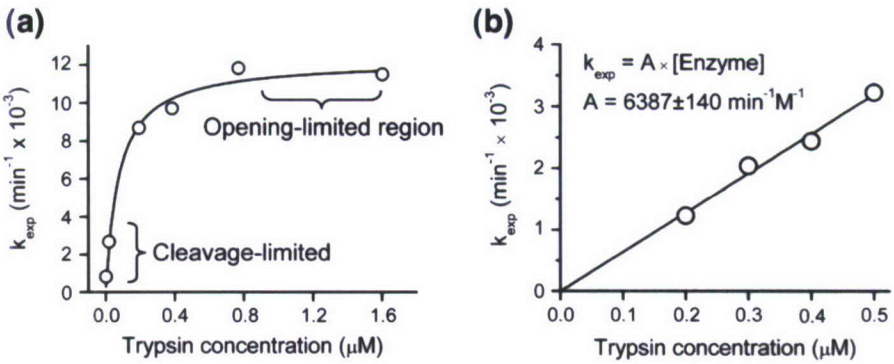


Figure 4. Determination of K_{op} and k_{open} from observed rate constants plotted *versus* trypsin concentration. (a) Over a wide range of concentrations the data can be described by a hyperbola. Under opening-limited conditions k_{exp} approaches a fixed limit: these plots are fit to the full kinetic equation (equation (7)), giving K_{op} and k_{open} values. Data shown is for the dimer at 37 °C. (b) In cleavage-limiting conditions, plots of k_{exp} are linear with respect to enzyme, with a slope of $K_{op} \cdot k_{cat}/K_M$. Data shown are for the capsid at 37 °C.

(k_{open}). Under these reaction conditions, the insensitivity of the resulting k_{open} value to the k_{cat}/K_M constant and enzyme concentration makes this an ideal initial screen for dynamic behavior. In order to determine K_{op}, reaction conditions are shifted to the other extreme, dilute trypsin, such that the trend between the experimental rate constant (k_{exp}) and enzyme concentration becomes linear ("cleavage-limited"; Figure 4(b)).

Unfolding rates and equilibria

Though k_{exp} varies between dimer and capsid by more than 20-fold, k_{open} is nearly identical between the two forms across the tested temperature range. The temperature dependence of kinetic parameters show a steady increase in the rate of the opening transition between 19 °C and 37 °C for both capsid and dimer forms of Cp149 (Figure 5(a)). In order to fully characterize K_{op} and k_{open} values for dimer and capsid forms of Cp149, digestion curves (Figure 4(a)) were generated for temperatures between 19 °C and 37 °C. When feasible, strict cleavage-limited reactions were also carried out to obtain a greater degree of precision (Figure 4(b)).

The uniform behavior between dimer and capsid with respect to opening rate is in contrast to the temperature dependence of the opening equilibrium: the equilibrium between open and closed states follows opposite trends for capsid and dimer. For capsid, K_{op} increases from 19 °C to 37 °C (the amount of open state changing from 0.05% to 1.24%; Figure 5(b)). In the dimer, at low temperature (19 °C) the region surrounding residues 127/128 is relatively exposed to enzymatic cleavage, with approximately 1.1% of the dimers existing in the open conformation. The equilibrium for dimer shifts toward the closed conformation with increasing temperature, reaching 0.259% at 37 °C. Since the opening rates for both capsid and dimer are nearly identical, these differences in exposure equilibrium can be attributed to the closing rate for the C-terminal region.

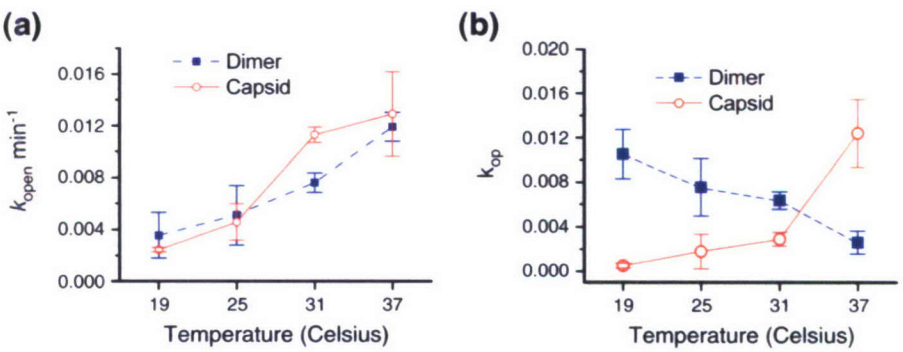


Figure 5. Kinetic and thermodynamic constants of Cp149 motion. (a) The rate of de-protection does not differ substantially between dimer and capsid forms of HBV. For both dimer and capsid a gradual increase in the rate of transition between protected and exposed states can be seen with increasing temperature. (b) Equilibrium between protected and cleavable states changes with temperature. Measured $K_{op} \cdot (k_{cat}/K_M)$ values are corrected for changes in intrinsic enzymatic activity for each temperature (Table 1), giving the equilibrium constant for de-protection. The capsid (open red circles, continuous line) exhibits the expected behavior with increasing exposure to enzymatic cleavage as the temperature increases. Unassembled dimer (filled blue squares, broken line) is seen to be most exposed to proteolysis at reduced temperatures, while increasing the temperature causes a shift in conformation to a more tightly folded state.

Thermodynamics

The equilibrium constant for the open–close transition, K_{op}, shows different behavior for dimer and capsid. Comparing dimer and capsid, values of K_{op} at a particular temperature differ by up to two orders of magnitude, and K_{op} trends between the two assembly states have opposite temperature dependence (Figure 5(b)). Van't Hoff plots of the temperature dependence of the closing transition give an enthalpy of -125.9 kJmol^{-1} and an entropic contribution of -368 J K^{-1}mol^{-1} for the closing process in capsids. For dimer K_{op} decreases with increasing temperature, indicating entropic stabilization of the folded state with $\Delta H = +54.5$ kJ mol^{-1} and $\Delta S = +224$ J K^{-1}mol^{-1} for closing. These values for dimer have the same trends as parameters calculated from capsid assembly experiments[23] and differ by less than an order of magnitude, suggesting a connection between folding of the C terminus and the capsid assembly reaction. Based upon the sign changes in enthalpy and entropy of the closing transition for dimer and capsid, the C terminus of Cp149 has fundamentally different behavior in the two assembly states.

These results show that dimer and capsid forms of Cp149, or the surrounding microenvironments, have heterogeneous dynamic behavior. At 37 °C, roughly 1% of capsid protein is in the open conformation, while dimer is almost entirely closed. At 19 °C, the properties are reversed. There is also no correlation between the folded state of the C terminus and the final assembly condition of Cp149: the open structural state observed for Cp149 can be achieved by both dimer and capsid under the appropriate conditions, while the primary difference between dimer and capsid is the opposite temperature dependence of the opening equilibrium.

Discussion

To date kinetic analysis of hydrolysis has been largely relegated to the study of model systems and small monomeric proteins. We have shown here that the energetics behind conformational flexibility, even within supramolecular complexes, can be measured using kinetic hydrolysis. Besides identification of the cleavage site using mass spectrometry, only standard lab equipment is required for these experiments. Preliminary SDS–PAGE and mass spectrometry of HBV Cp149 reaction products showed that proteolysis in both dimer and capsid forms exclusively targets the trypsin site R127/T128. Because the initial proteolytic reaction is focused on a single site, the kinetic analysis is practical and unambiguous. Computational modeling of proteolytic reactions indicates that even a minor secondary cleavage pathway can have a significant impact on the calculated equilibrium and rates.[63,64] These concerns make it vital that measurements of dynamic motion based on the disappearance of starting material be the result of a single cleavage pathway. Previously this requirement has been met by limiting the application of kinetic hydrolysis to small proteins that are well folded.[61,62,65] Supramolecular complexes can be amenable to the same analyses, provided suitable precautions are met.

The application of kinetic hydrolysis to supramolecular complexes introduces a series of structural concerns. It is possible that structural rearrangements such as dimer dissociation, capsid integrity, or destabilizing cleavages could all affect the measured rates, but these factors were found to be absent under the conditions studied. Capsids are very robust and no detectable dissociation has been observed with the buffer conditions used in these experiments.[23] At 37 °C, dimer is seen to have a reduced K_{op} value compared to capsid, ruling out dimer dissociation as an intermediate in the reaction pathway of capsid proteolysis. Dynamic light scattering and size-exclusion chromatography experiments show no general capsid destabilization during the course of the reactions, and by collecting data only on solutions with greater than 70% completely intact protein, destabilization is further minimized. Localized effects as a result of cleavage within the clusters of five or six C termini at the capsid pores cannot be ruled out, but the loss of the cleavage peptide at the terminus is not expected to reduce steric protection for neighboring dimers. Besides accessibility, the cleavage site stability is maintained as the reaction progresses. The stabilizing hydrophobic cluster consisting of amino acids L140, F110, and Y118 is formed within one dimer, and by restricting data collection to only 30% completion, there is less than 10% likelihood of a second cleavage occurring adjacent to an already cleaved site.

We find that HBV capsids and core protein dimers are very dynamic compared to the small protein examples known. The exposure of the R127/T128 cleavage site is not a rare process, with equilibrium values ranging between 5×10^{-4} and 0.0124. Similar studies of other proteins have found values more than 200-fold lower.[50,61,66] It may also be possible that the magnitude of motion required for transition between the folded and protease-accessible states is far greater than that required for biological activity in the native particles, resulting in effectively higher exposure for native interactions. While the measured K_{op} equilibrium is larger than values observed for other systems, k_{open} trends are relatively similar.[50] This indicates that the high exposure distribution of the C terminus is maintained by a very slow closing rate. In a biological context, these parameters give an average of three individual monomers per capsid in the open state. Because all reactions are carried out in solution and represent the average for $\sim 2 \times 10^{12}$ individual capsids, it is impossible to say whether the opening transition is independent, or cooperative in the context of a capsid. It is attractive to speculate that several adjacent subunits around a 5-fold or quasi-6-fold vertex open at the same time.

Similar to the equilibrium trends, the exposure rates of the cleavage site indicate the dynamic nature of Cp149. The slowest opening transition to a cleavage-accessible state is found for capsid at 19 °C, with a rate constant of 3.5 day^{-1}. While this is comparable to values seen for sets of stable amide protons in HDX experiments,[67–69] the motions required for the "open" state in hydrogen-exchange are extremely small, sometimes less than an angstrom unit.[52] In contrast, proteolysis requires a substantial region of unfolded backbone,[44] and in the case of Cp149, steric constraints from the four-helix turrets increase the required exposure for cleavage. The total translation, between the conformation in the crystal structure to an estimated cleavage-accessible unfolded state, may necessitate as much as a 13 Å motion for the R127/T128 pair alone. This large-amplitude motion is sampled on a moderate timescale for even the most strongly folded conditions studied (19 °C, capsid).

The Cp149 C terminus likely populates an ensemble of conformations in solution. The application of kinetic hydrolysis can only resolve the rate of transition across an arbitrary boundary between protease-protected and accessible conformations. A two-state approximation has been a sufficient and successful model for kinetic hydrolysis studies of other proteins.[60,62,65,70,71] Our results reaffirm applicability of a two-state closed/open system to HBV Cp149 dimer and capsid. While there may be subtle complications beyond the two-state behavior, any deviations present are insignificant within the error of the data. The kinetics follow the predicted pattern for a partially protected peptide backbone: cleavage rates of the dimer are approximately 150-fold slower compared to the unstructured model peptide. However, the observed opening rates (Figure 5(a)) far exceed predicted values for non-dynamic protein structures.[52] The observed rate of digestion also reaches an upper limit in the opening-limited conditions of high enzyme concentration (Figure 4(a)), indicating that transient interconversion is responsible for the protection effect. This phenomenon has been observed with other known two-state systems.[50,70]

Structures containing atomic-level detail of the C terminus of Cp149 in capsid form, crystallized at ~21 °C to 24 °C, are available.[13,14] These models are in good agreement that residues G112–T142 adopt a defined, folded helix-loop-arm conformation, with L140 wrapping around to meet with residues F110 and Y118 (see Figures 1(b) and 6(a) and (c)). The site of cleavage, R127/T128, is located near the terminus of helix 5 in this helix-loop-arm. The relative size of the protease is shown next to the capsid in Figure 1. Rigid body docking of the crystal structures indicates that substantial unfolding, as described above, must take place to allow the cleavage site to reach the enzyme active site. In this case, the static models from crystallography and cryo-electron microscopy are silent to the observed dynamics.

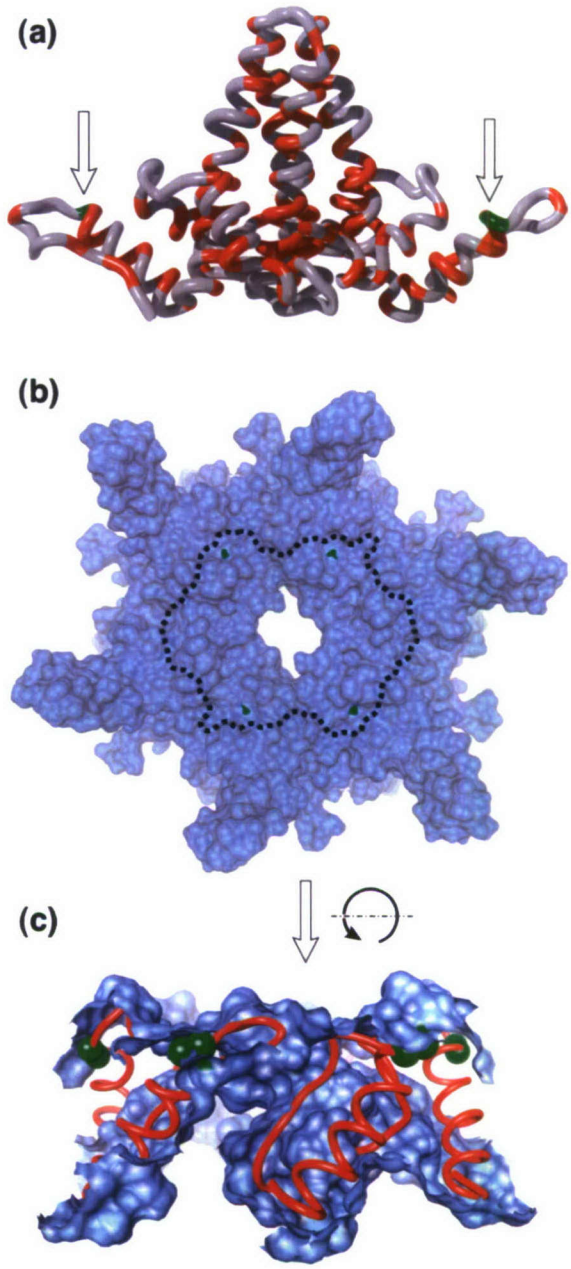


Figure 6. Solution phase dynamics localize to a well ordered helix in the structural model. The structure of HBV Cp149 is shown in dimer and capsid forms ((a) and (b) and (c), respectively: not to scale). (a) Dimer: the trypsin cleavage site between residues 127 and 128 is indicated with arrows. Despite the presence of alternative protease sites for trypsin and thermolysin (shown in red), cleavage is restricted to the end of helix 5 (in green). (b) Local overhead view of a quasi-6-fold center of symmetry within the assembled capsid. Shown as a solvent-excluded surface, the cleavage sites (green) are barely exposed to solvent. The observed cleavage location (green) is central to the packing interface between dimers at the 5-fold and 6-fold axes of symmetry in the capsid. The indicated boundary shows the extent of Figure 5(c). (c) Cutaway of Figure 5(b), rotated 90° away from the viewer. Dimers make close contact centered at helix 5 (red chain), preventing access to the cleaved bond (green spheres).

Based on the differences between structural data and our dynamics results, we postulate that free dimer adopts a conformation where the C-terminal domain is loosely held in place by hydrophobic interactions in a manner that is distinct from capsid. In capsids, the cleavage site is dynamic but leaves most hydrophobic contacts intact during breathing modes, which results in the change of sign for the thermodynamic parameters. This model is also consistent with previous studies that have detailed the thermodynamics of capsid formation.[23] The temperature dependence of dimer opening also clarifies the role of small molecules such as HAP and Zn^{2+} in the allosteric regulation of assembly.[14,21,27] At 21 °C Zn^{2+} induces trapped assembly while at 37 °C it leads to virus-like particles. The HAP molecules strongly enhance assembly and have been shown to pack against the C-terminal helix and arm.[14] In the assembled capsid the C-terminal arm is tightly packed into the inter-dimer hydrophobic interfaces.[13,14] Our results indicate that the assembled HBV capsid can be described as a thermodynamic balancing act between the inter-subunit affinity, which promotes assembly and localized opening which allow the C-terminal domain to be dynamic.

The wealth of data available on the behavior of the HBV capsid allows these results to be put into context with respect to the biology of the virus. *In vivo*, the dynamic behavior of the C terminus plays a general role in the lifecycle of HBV.[72] The construct used for these experiments, with deletion of the C-terminal arginine-rich region, may be thought of as the minimal structure which still retains the features of the native protein.[73] By using this variant, which is well studied from the perspective of both assembly and structure, the kinetic hydrolysis measurements can be compared to data, which is unavailable for the full-length capsid protein. However, the full-length protein also has behavior which is in agreement with our observations of Cp149. The RNA binding domain, the C terminus (residues 150–183) of the full-length core protein, is located on the interior of the capsid and directly exposed to the nucleic acid. However, the phosphorylation state is altered during assembly, maturation,[32,37,41,42,74,75] and nuclear localization of assembled particles.[33,35,36,76] Transient exposure of this domain will allow reaction with kinases, phosphatases, and the nuclear transport machinery. Secretion of mature cores from the cell must take advantage of other host transport machinery, probably including the ubiquitin ligase, Nedd4.[38] However, Nedd4 is predicted to bind to a late domain PPAY (residues 129–132; Figure 1(c)) that is largely buried.[38] Our study indicates that PPAY, which is adjacent to the cleavage site, will be transiently exposed. We can reasonably hypothesize that the nucleic acid content of the core will affect the breathing modes that expose the RNA-binding and late domain sequences, providing a mechanism for DNA synthesis to signal the next steps in the virus life cycle.

Conclusion

The use of proteolytic digestion to probe dynamic behavior of proteins is well established and provides an alternative approach to spectroscopic means for obtaining thermodynamic information.[45,50,61,62] However, the advantages it provides (simplicity despite structural size and complexity; amenability to varying buffer, pH, and temperature conditions; and sensitivity to large-scale motion) have not been fully explored. We have demonstrated the utility of a quantitative kinetic proteolysis approach to study the 4 MDa hepatitis B viral capsid. Our findings indicate that the C terminus is highly dynamic, resulting in transient exposure to the surface of the capsid. Similarly internalized domains in other icosahedral capsids are known to be dynamic and exposed in solution.[7,49,55,77,78] The detailed characterization of this motion for the HBV capsid has shown that the dynamics are highly temperature dependent. In dimer, the opening transition appears to have negative entropy and positive enthalpy, while in capsid the reverse is true. This thermodynamic switch between oligomeric states may help explain the variety of roles observed for the C terminus, where structural localization is predicted to vary in the course of the HBV lifecycle. The approximate match between energetics of C terminus "closing" in the dimer and capsid assembly may indicate that folding in this region is part of the pathway for capsid assembly, although further investigation is needed to clarify the role of any intermediates. The measurements of kinetic and thermodynamic parameters made possible with kinetic hydrolysis provide the first mechanistic evidence for dynamic structural behavior within the highly biologically relevant region of the hepatitis B core protein.

Materials and Methods

HBV protein preparation

HBV capsids were produced from recombinant $Cp149_2$, expressed in BL21 DE3 *Escherichia coli* and purified with multiple cycles of size-exclusion chromatography as previously described.[79] Capsids were stored in 50 mM Hepes (pH 7.5), 0.5 M NaCl, 2 mM DTT at 4 °C. Purified capsids were dissociated with 3 M urea on ice and isolated with size-exclusion chromatography to produce a dimer stock. Storage conditions for the dimer were 50 mM Na_2CO_3 (pH 9.6), 2 mM DTT at 4 °C. The stock solutions of Cp149 were regularly checked using FPLC size-exclusion chromatography (100 cm Sephacryl 500; GE Healthcare Inc.; Chalfont St. Giles, GBR) with detection at 280 nm and dynamic light scattering Brookhaven 90Plus (Brookhaven Instruments Corp.; Holtsville, NY).

Immediately before use, samples were diluted into reaction buffer. Although previous research has shown the robust stability of Cp149 dimers and capsid under the buffer conditions used,[17] example reactions were run within a temperature-controlled Brookhaven 90Plus dynamic light scattering (DLS) instrument to monitor the

assembly state during the progress of digestion. DLS was measured at 90° using a 661 diode laser, and correlation functions were fit using a non-negatively constrained least-squares analysis. As the proteolysis progressed, no decrease in the primary signal was observed, nor were there any smaller particles or aggregate material seen. Capsid integrity was also evaluated by size exclusion chromatography with a Superose 6 column (Amersham-Pharmacia, Piscataway, NJ). Buffer conditions were identical to those used for proteolysis reactions (see below), and detection was followed at 280 nm. The experimental results provided additional evidence for the continuing integrity of the capsids (see Results).

Proteolysis reactions

Reaction solutions consisted of 0.5 mg/ml of HBV protein, 150 mM NaCl, 100 mM Tris (pH 7.5), and variable concentrations of sequencing-grade modified trypsin (Promega Corp.; Fitchburg, WI). All buffers were individually prepared to account for the temperature sensitivity of Tris. Trypsin was stored in frozen aliquots in 1 mM HCl: any dilutions were made immediately before use. Each reaction was initiated by addition of an appropriate volume of trypsin and vortexed. A typical experiment consisted of 50–100 μl final volume incubated in a thermocycler for the duration of the digestion. Aliquots of 3 μl were removed in triplicate from the reaction and immediately diluted into 12 μl of gel loading buffer, vortexed, and heated to 100 °C for 5 min. Following boiling, samples were cooled on ice, briefly centrifuged to collect condensation, vortexed, and stored on ice until use. Prior to gel loading, the frozen samples were simultaneously boiled at 100 °C for 5 min, cooled to 25 °C, centrifuged, vortexed, and maintained at 25 °C during gel loading. These additional precautions for sample preparation were used to maximize reproducibility in gel loading. Samples were run on 4%–20% (w/v) gradient and 18% non-gradient SDS–polyacrylamide gels (Bio-Rad, Hercules, CA) with Tris-glycine running buffer. Running conditions were 200 V and the minimal time required to ensure separation between intact protein and cleaved products, typically 25 min. Gels were prepared according to manufacturer's instructions for staining with colloidal Coomassie or Sypro Ruby (Pierce Biotechnology, Rockford, IL), and scanned with a Typhoon 9410 laser scanner (GE Healthcare Inc.) for Sypro stain and Epson 4180 scanner (Seiko Epson Co, Suwa, JPN) for Coomassie stain.

Preliminary experiments were conducted to ensure that the subset of the hydrolysis progression assayed for quantitative measurements was representative of the overall trend and did not span any discontinuities (Figure 2(a) and (c)). To ensure the integrity of the viral capsids, for quantitative reactions all data were collected from time points containing greater than 70% uncleaved protein, as determined by gel densitometry. For each enzyme concentration, sampling times were adjusted as necessary for best coverage of this region. Samples were removed from reactions either as individual time stamps (Figure 2(a) and (c)) or in triplicate at each time point (Figure 2(b) and (d)). The latter approach allowed an evaluation of sampling and quantitation error.

Densitometry

While it is necessary to limit the extent of digestion when working with a protein complex to insure integrity, this complicates densitometry because the observed differences in band intensities are limited and subtle. By restricting proteolytic cleavage to less than 30% completion, the change in intact protein intensity represents a very small variable range for typical densitometry. In order to maximize the accuracy of the quantitation, several novel methods were employed. Image acquisition is the first step of the processing pipeline which can have an impact on the final result: concerns typically include linearity, sensitivity, and dynamic range.[80] However, for quantitation of band intensities, which are never more than ±15%, these effects are minimal. In order to test this assumption, gel images were stained with the standard fluorescent stain SYPRO ruby and image acquisition was performed using a Typhoon 9400 laser scanner. Gels were then stained with the colloidal Coomassie stain and imaged with an Epson 4180 flatbed CCD scanner. For the types of gel images used, comparative analysis showed no significant deviation between the two methods. Since acquisition introduces minimal error for homogeneous images, detection and quantitation of gel bands was addressed as the primary source of variability. Image manipulation was performed using lossless operations in Adobe Photoshop CS, followed by analysis using ImageJ (Rasband, U S National Institutes of Health† with a manual watershed isolation method and a high-resolution modified version of the GelAnalyzer module. Band intensities were plotted and fit in Origin 7.0 to the first-order equation:

$$I_t = I_0 \cdot e^{-k_{exp}t} \tag{2}$$

where I_t is the measured intact protein at time t, I_0 is the intensity at $t=0$, and k_{exp} is a first-order rate constant. These precautions made it possible to quantify band intensities to an acceptable degree of precision: the total variability for replicate lanes was typically less than 4% relative standard deviation.

Mass spectrometry

Proteolysis reactions for mass spectrometry analysis were conducted identical to those for gel-based analysis except for termination; in some situations one reaction was simultaneously analyzed with gels and mass spectrometry. Quenching of proteolysis was performed *via* multiple methods to avoid biasing peptide detection. These included dropping the pH, addition of irreversible trypsin inhibitors such as PMSF, boiling, spotting directly on to a MALDI plate, and direct LCMS injection. For electrospray-based MS analysis sample quantity was varied over four orders of magnitude spanning the typical detection range. Electrospray instrumentation included an Agilent XCT Plus (Agilent Technologies, Santa Clara, CA), Bruker micrOTOF LC (Bruker Daltonik GmbH, Bremen, DEU), and Waters QToF Premier (Waters Corp., Milford, MA), all interfaced with reverse-phase HPLC (C8 or C18: Phenomenex Inc., Torrance, CA). MALDI analysis was performed with a variety of sample/matrix ratios and both sinapinic acid and α-Cyano-4-hydroxycinnamic acid, on a Bruker BiFlex III.

Intrinsic rate determination of trypsin cleavage

To determine the activity of trypsin upon the observed cleavage sequence in absence of any structural or steric

† http://rsb.info.nih.gov/ij/

constraints, the fluorogenic peptide Abz-IRTPP-Tyr(3-NO_2) was characterized as a trypsin substrate. While the peptide is intact, the FRET fluorescence-quenching nitrotyrosine residue decreases intrinsic fluorescence of the 2-aminobenzamide (Abz) at the C terminus. Upon cleavage at the R/T trypsin site, the Abz fluorescence is restored.[81] The synthetic peptide was purchase from New England Peptide (Gardner, MA) at >99% pure and used without further purification. Assays were carried out for each temperature under the same conditions as for proteolytic reactions: 150 mM NaCl and 100 mM Tris (pH 7.5). Peptide concentration was varied from 150 nM to 75 μM in order to generate K_M curves. Buffered solutions containing the peptide were pre-incubated in a temperature-controlled and stirred cuvette with a PerkinElmer LS50B fluorometer (PerkinElmer, Waltham, MA) and reactions were initiated by addition of trypsin to a final concentration of 9.6 nM to 385 nM. Changes in Abz fluorescence were observed with an excitation wavelength of 318 nm and an emission wavelength of 420 nm. Depending on trypsin concentration, analysis was carried out by fitting data to either initial-velocity Michaelis–Menten kinetic models or with an exponential-rise model:[82]

$$F_t = F_0 + (F_{max} - F_0) \cdot e^{(-V/K \cdot t)} \quad (3)$$

Where F_t is the measured fluorescence, F_0 is the fluorescence at $t=0$, F_{max} is fluorescence at $t=\infty$, and V/K is a pseudo-first-order rate constant. From this exponential-rise model k_{cat}/K_M was determined by dividing V/K by the enzyme concentration of the reaction.

Kinetic models of proteolysis

For protein domains that partition between folded and unfolded conformations, the rate of cleavage can be modeled with a two-state equilibrium system:

$$Closed \underset{k_{close}}{\overset{k_{open}}{\rightleftharpoons}} Open \xrightarrow{k_{cleave}} Cleaved \quad (1)$$

Under normal conditions without the presence of protease, an equilibrium is established between open and closed forms of the protein, giving:

$$K_{op} = \frac{k_{open}}{k_{close}} \quad (4)$$

and with: $[Closed] + [Open] = [Total]$:

$$[Open] = \frac{K_{op}}{K_{op} + 1} \cdot [Total] \quad (5)$$

For typical cases with small values of K_{op}, equation (5) simplifies to $[Open] = K_{op} \cdot [Total]$, removing the denominator term. This simplification can be propagated through derivative equations without excessive error. In the presence of a proteolytic enzyme, the intrinsic cleavage rate constant k_{cleave} can be calculated as:

$$k_{cleave} = \frac{k_{cat}}{K_M} \cdot [E] \quad (6)$$

with k_{cat}/K_M being the standard catalytic constants for trypsin cleavage of the particular peptide sequence and $[E]$ is the concentration of trypsin. The pseudo-first-order simplification of the Michaelis–Menten equation is justified because the measured K_M value for this sequence was 20 μM, which is much greater than the concentration of available HBV cleavage sites. A steady-state simplification of the system in equation (1) can be used to approximate the observed first-order rate constant k_{exp} as a function of the component constants:

$$k_{exp} = \frac{k_{open} k_{cleave}}{k_{close} + k_{cleave}} \quad (7)$$

This produces a hyperbolic trend when sufficient quantities of enzyme are used in the digestions. When $k_{cat}/K_M \cdot [E] \gg k_{close}$, any protein in the open conformation is immediately cleaved and k_{exp} approaches a maximum limit. This is termed "opening-limited conditions", and under such situations $k_{exp} = k_{open}$. Since in practice it is rarely possible to achieve perfectly enzyme-saturated levels, k_{open} is typically extrapolated by fitting equation (7) with non-linear regression to data collected across a wide range of enzyme concentrations (Figure 3(a)).

Alternatively, when enzyme concentrations are low enough that cleavage is slow relative to the equilibrium process ($k_{cat}/K_M \cdot E \ll k_{close}$), the concentration of protein in the open conformation remains a function of the remaining total uncleaved protein (equation (5)). By combining equations (5) and (6), the linear relationship between enzyme and experimental rate constant:

$$k_{exp} = \frac{k_{cat}}{K_M} \cdot [E] \cdot K_{op} \quad (8)$$

can be modeled for reactions with low concentration of enzyme. This is termed "cleavage-limited conditions", and data from such reactions can be fit with a linear regression (Figure 3(b)).

Once K_{op} values have been obtained, $\Delta G_{deprotection}$ is calculated as $-RT \cdot \ln(K_{op})$. Not to be confused with the energetics of proteolytic cleavage, this thermodynamic parameter is the free energy difference between the open and closed conformation states of the protein.

Acknowledgements

The authors acknowledge the technical support of Pablo Ceres, Angela Lee, and Jennifer M Johnson and thank Richard Kriwacki, Kimberly Hilmer, and Jason Lanman for critical reading of the manuscript. A.Z. was supported by grant R56 AI068883. B.B. acknowledges that the project was supported by NIH grant number P20 RR-020185 from the COBRE Program of the National Center For Research Resources. B.B. also received support from the Center for Bio-Inspired Nanomaterials (Office of Naval Research grant N00014-06-01-1016). The Murdock Charitable Trust supports the Mass Spectrometry Facility at MSU.

References

1. Wales, D. J. (2005). The energy landscape as a unifying theme in molecular science. *Phil. Trans.ser. A Math. Phys. Eng. Sci.* **363**, 357–375.
2. Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A.,

Skalicky, J. J., Kay, L. E. & Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature*, **438**, 117–121.
3. Vendruscolo, M. (2007). Determination of conformationally heterogeneous states of proteins. *Curr. Opin. Struct. Biol.* **17**, 15–20.
4. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M. & Cate, J. H. (2005). Structures of the bacterial ribosome at 3.5 Å resolution. *Science*, **310**, 827–834.
5. Frey, S., Richter, R. P. & Gorlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science*, **314**, 815–817.
6. Witz, J. & Brown, F. (2001). Structural dynamics, an intrinsic property of viral capsids. *Arch. Virol.* **146**, 2263–2274.
7. Johnson, J. E. (2003). Virus particle dynamics. *Adv. Protein Chem.* **64**, 197–218.
8. Cremeens, M. E., Fujisaki, H., Zhang, Y., Zimmermann, J., Sagle, L. B., Matsuda, S. *et al.* (2006). Efforts toward developing direct probes of protein dynamics. *J. Am. Chem. Soc.* **128**, 6028–6029.
9. Jarymowycz, V. A. & Stone, M. J. (2006). Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chem. Rev.* **106**, 1624–1671.
10. Tehei, M., Daniel, R. & Zaccai, G. (2006). Fundamental and biotechnological applications of neutron scattering measurements for macromolecular dynamics. *Eur. Biophys. J.* **35**, 551–558.
11. Stannard, L. M. & Hodgkiss, M. (1979). Morphological irregularities in Dane particle cores. *J. Gen. Virol.* **45**, 509–514.
12. Crowther, R. A., Kiselev, N. A., Bottcher, B., Berriman, J. A., Borisova, G. P., Ose, V. & Pumpens, P. (1994). Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell*, **77**, 943–950.
13. Wynne, S. A., Crowther, R. A. & Leslie, A. G. (1999). The crystal structure of the human hepatitis B virus capsid. *Mol. Cell.* **3**, 771–780.
14. Bourne, C. R., Finn, M. & Zlotnick, A. (2006). Global structural changes in hepatitis B virus capsids induced by the assembly effector HAP1. *J. Virol.* **80**, 11055–11061.
15. Dryden, K. A., Wieland, S. F., Whitten-Bauer, C., Gerin, J. L., Chisari, F. V. & Yeager, M. (2006). Native hepatitis B virions and capsids visualized by electron cryomicroscopy. *Mol. Cell.* **22**, 843–850.
16. Zhou, S. & Standring, D. N. (1992). Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. *Proc. Natl Acad. Sci. USA*, **89**, 10046–10050.
17. Singh, S. & Zlotnick, A. (2003). Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly. *J. Biol. Chem.* **278**, 18249–18255.
18. Bottcher, B., Wynne, S. A. & Crowther, R. A. (1997). Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature*, **386**, 88–91.
19. Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J. & Steven, A. C. (1997). Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature*, **386**, 91–94.
20. Wingfield, P. T., Stahl, S. J., Williams, R. W. & Steven, A. C. (1995). Hepatitis core antigen produced in *Escherichia coli*: subunit composition, conformational analysis, and *in vitro* capsid assembly. *Biochemistry*, **34**, 4919–4932.
21. Stray, S. J., Ceres, P. & Zlotnick, A. (2004). Zinc ions trigger conformational change and oligomerization of hepatitis B virus capsid protein. *Biochemistry*, **43**, 9989–9998.
22. Ceres, P., Stray, S. J. & Zlotnick, A. (2004). Hepatitis B virus capsid assembly is enhanced by naturally occurring mutation F97L. *J. Virol.* **78**, 9538–9543.
23. Ceres, P. & Zlotnick, A. (2002). Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry*, **41**, 11525–11531.
24. Zlotnick, A., Johnson, J. M., Wingfield, P. W., Stahl, S. J. & Endres, D. (1999). A theoretical model successfully identifies features of hepatitis B virus capsid assembly. *Biochemistry*, **38**, 14644–14652.
25. Choi, Y., Gyoo Park, S., Yoo, J. H. & Jung, G. (2005). Calcium ions affect the hepatitis B virus core assembly. *Virology*, **332**, 454–463.
26. Stray, S. J. & Zlotnick, A. (2006). BAY 41-4109 has multiple effects on Hepatitis B virus capsid assembly. *J. Mol. Recognit.*.
27. Stray, S. J., Bourne, C. R., Punna, S., Lewis, W. G., Finn, M. G. & Zlotnick, A. (2005). A heteroaryldihydropyrimidine activates and can misdirect hepatitis B virus capsid assembly. *Proc. Natl Acad. Sci. USA*, **102**, 8138–8143.
28. Zlotnick, A., Ceres, P., Singh, S. & Johnson, J. M. (2002). A small molecule inhibits and misdirects assembly of hepatitis B virus capsids. *J. Virol.* **76**, 4848–4854.
29. Hacker, H. J., Deres, K., Mildenberger, M. & Schroder, C. H. (2003). Antivirals interacting with hepatitis B virus core protein and core mutations may misdirect capsid assembly in a similar fashion. *Biochem. Pharmacol.* **66**, 2273–2279.
30. Deres, K., Schroder, C. H., Paessens, A., Goldmann, S., Hacker, H. J., Weber, O. *et al.* (2003). Inhibition of hepatitis B virus replication by drug-induced depletion of nucleocapsids. *Science*, **299**, 893–896.
31. Butz, K., Denk, C., Fitscher, B., Crnkovic-Mertens, I., Ullmann, A., Schroder, C. H. & Hoppe-Seyler, F. (2001). Peptide aptamers targeting the hepatitis B virus core protein: a new class of molecules with antiviral activity. *Oncogene*, **20**, 6579–6586.
32. Yu, M. & Summers, J. (1994). Phosphorylation of the duck hepatitis B virus capsid protein associated with conformational changes in the C terminus. *J. Virol.* **68**, 2965–2969.
33. Rabe, B., Vlachou, A., Pante, N., Helenius, A. & Kann, M. (2003). Nuclear import of hepatitis B virus capsids and release of the viral genome. *Proc. Natl Acad. Sci. USA*, **100**, 9849–9854.
34. Kock, J., Kann, M., Putz, G., Blum, H. E. & Von Weizsacker, F. (2003). Central role of a serine phosphorylation site within duck hepatitis B virus core protein for capsid trafficking and genome release. *J. Biol. Chem.* **278**, 28123–28129.
35. Yeh, C. T., Liaw, Y. F. & Ou, J. H. (1990). The arginine-rich domain of hepatitis B virus precore and core proteins contains a signal for nuclear transport. *J. Virol.* **64**, 6141–6147.
36. Kann, M., Sodeik, B., Vlachou, A., Gerlich, W. H. & Helenius, A. (1999). Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. *J. Cell Biol.* **145**, 45–55.
37. Perlman, D. H., Berg, E. A., O'Connor P, B., Costello, C. E. & Hu, J. (2005). Reverse transcription-associated dephosphorylation of hepadnavirus nucleocapsids. *Proc. Natl Acad. Sci. USA*, **102**, 9020–9025.
38. Rost, M., Mann, S., Lambert, C., Doring, T., Thome, N.

& Prange, R. (2006). Gamma-adaptin, a novel ubiquitin-interacting adaptor, and Nedd4 ubiquitin ligase control hepatitis B virus maturation. *J. Biol. Chem.* **281**, 29297–29308.
39. Kann, M. & Gerlich, W. H. (1994). Effect of core protein phosphorylation by protein kinase C on encapsidation of RNA within core particles of hepatitis B virus. *J. Virol.* **68**, 7993–8000.
40. Lan, Y. T., Li, J., Liao, W. & Ou, J. (1999). Roles of the three major phosphorylation sites of hepatitis B virus core protein in viral replication. *Virology*, **259**, 342–348.
41. Kock, J., Nassal, M., Deres, K., Blum, H. E. & von Weizsacker, F. (2004). Hepatitis B virus nucleocapsids formed by carboxy-terminally mutated core proteins contain spliced viral genomes but lack full-size DNA. *J. Virol.* **78**, 13812–13818.
42. Le Pogam, S., Chua, P. K., Newman, M. & Shih, C. (2005). Exposure of RNA templates and encapsidation of spliced viral RNA are influenced by the arginine-rich domain of human hepatitis B virus core antigen (HBcAg 165–173). *J. Virol.* **79**, 1871–1887.
43. Matthyssens, G. E., Simons, G. & Kanarek, L. (1972). Study of the thermal-denaturation mechanism of hen egg-white lysozyme through proteolytic degradation. *Eur. J. Biochem.* **26**, 449–454.
44. Hubbard, S. J., Campbell, S. F. & Thornton, J. M. (1991). Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. *J. Mol. Biol.* **220**, 507–530.
45. Carey, J. (2000). A systematic and general proteolytic method for defining structural and functional domains of proteins. *Methods Enzymol.* **328**, 499–514.
46. Imoto, T., Fukuda, K. & Yagishita, K. (1974). A study of the native-denatured (N-D) transition in lysozyme I. Detection of the transition by product analyses of protease digests. *Biochim. Biophys. Acta*, **336**, 264–269.
47. Polverino de Laureto, P., Scaramella, E., Frigo, M., Wondrich, F. G., De Filippis, V., Zambonin, M. & Fontana, A. (1999). Limited proteolysis of bovine alpha-lactalbumin: isolation and characterization of protein domains. *Protein Sci.* **8**, 2290–2303.
48. Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B. & Wetzel, R. (2001). Structural features of the Abeta amyloid fibril elucidated by limited proteolysis. *Biochemistry*, ,**40**, 11757–11767.
49. Bothner, B., Dong, X. F., Bibbs, L., Johnson, J. E. & Siuzdak, G. (1998). Evidence of viral capsid dynamics using limited proteolysis and mass spectrometry. *J. Biol. Chem.* **273**, 673–676.
50. Park, C. & Marqusee, S. (2004). Probing the high energy states in proteins by proteolysis. *J. Mol. Biol.* **343**, 1467–1476.
51. Konermann, L. & Simmons, D. A. (2003). Protein-folding kinetics and mechanisms studied by pulse-labeling and mass spectrometry. *Mass Spectrom. Rev.* **22**, 1–26.
52. Wales, T. E. & Engen, J. R. (2006). Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass. Spectrom. Rev.* **25**, 158–170.
53. Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I. & Wright, P. E. (1996). Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl Acad. Sci. USA*, **93**, 11504–11509.
54. Fricks, C. E. & Hogle, J. M. (1990). Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. *J. Virol.* **64**, 1934–1945.
55. Bothner, B., Taylor, D., Jun, B., Lee, K. K., Siuzdak, G., Schultz, C. P. & Johnson, J. E. (2005). Maturation of a tetravirus capsid alters the dynamic properties and creates a metastable complex. *Virology*, **334**, 17–27.
56. Speir, J. A., Bothner, B., Qu, C., Willits, D. A., Young, M. J. & Johnson, J. E. (2006). Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics. *J. Virol.* **80**, 3582–3591.
57. Hogle, J. M. (2002). Poliovirus cell entry: common structural themes in viral cell entry pathways. *Annu. Rev. Microbiol.* **56**, 677–702.
58. Havlis, J., Thomas, H., Sebela, M. & Shevchenko, A. (2003). Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal. Chem.* **75**, 1300–1306.
59. Finehout, E. J., Cantor, J. R. & Lee, K. H. (2005). Kinetic characterization of sequencing grade modified trypsin. *Proteomics*, **5**, 2319–2321.
60. Prudova, A., Bauman, Z., Braun, A., Vitvitsky, V., Lu, S. C. & Banerjee, R. (2006). S-Adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity. *Proc. Natl Acad. Sci. USA*, **103**, 6489–6494.
61. Wildes, D., Anderson, L. M., Sabogal, A. & Marqusee, S. (2006). Native state energetics of the Src SH2 domain: evidence for a partially structured state in the denatured ensemble. *Protein Sci.* **15**, 1769–1779.
62. Arnold, U., Köditz, J., Markert, Y. & Ulbrich-Hofmann, R. (2005). Local fluctuations vs global unfolding of proteins investigated by limited proteolysis. *Biocatal. Biotransform.* **23**, 159–167.
63. Vorob'ev, M. M., Dalgalarrondo, M., Chobert, J. M. & Haertle, T. (2000). Kinetics of beta-casein hydrolysis by wild-type and engineered trypsin. *Biopolymers*, **54**, 355–364.
64. Schellenberger, V., Siegel, R. A. & Rutter, W. J. (1993). Analysis of enzyme specificity by multiple substrate kinetics. *Biochemistry*, **32**, 4344–4348.
65. Park, C. & Marqusee, S. (2005). Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. *Nature Methods*, **2**, 207–212.
66. Young, T. A., Skordalakes, E. & Marqusee, S. (2007). Comparison of proteolytic susceptibility in phosphoglycerate kinases from yeast and *E. coli*: modulation of conformational ensembles without altering structure or stability. *J. Mol. Biol.* **368**, 1438–1447.
67. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 75–86.
68. Sari, N., Ruan, B., Fisher, K. E., Alexander, P. A., Orban, J. & Bryan, P. N. (2007). Hydrogen-deuterium exchange in free and prodomain-complexed subtilisin. *Biochemistry*, **46**, 652–658.
69. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1994). Protein stability parameters measured by hydrogen exchange. *Proteins: Struct. Funct. Genet.* **20**, 4–14.
70. Imoto, T., Yamada, H. & Ueda, T. (1986). Unfolding rates of globular proteins determined by kinetics of proteolysis. *J. Mol. Biol.* **190**, 647–649.
71. Fontana, A., de Laureto, P. P., Spolaore, B., Frare, E., Picotti, P. & Zambonin, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* **51**, 299–321.
72. Knipe, D. M., Howley, P. M., Griffin, D. E., Lamb, R. A. & Martin, M. A. (2006). *Fields Virology* 5th edit., 2 vols, Lippincott Williams & WilkinsHagerstown, MD.

73. Zlotnick, A., Cheng, N., Conway, J. F., Booy, F. P., Steven, A. C., Stahl, S. J. & Wingfield, P. T. (1996). Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. *Biochemistry*, **35**, 7412–7421.
74. Ganem, D. & Schneider, R. J. (2001). Hepadnaviridae: The viruses and their replication. In *Fields Virology* (Knipe, D. M. & Howley, P. M., eds), 4th edit. Lippincott, Williams & Wilkins, Philadelphia.
75. Nassal, M. (1992). The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J. Virol.* **66**, 4107–4116.
76. Yeh, C. T., Hong, L. H., Ou, J. H., Chu, C. M. & Liaw, Y. F. (1996). Characterization of nuclear localization of a hepatitis B virus precore protein derivative P22. *Arch. Virol.* **141**, 425–438.
77. Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. (1995). Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure*, **3**, 63–78.
78. Lewis, J. K., Bothner, B., Smith, T. J. & Siuzdak, G. (1998). Antiviral agent blocks breathing of the common cold virus. *Proc. Natl Acad. Sci. USA*, **95**, 6774–6778.
79. Zlotnick, A., Lee, A., Bourne, C. R., Johnson, J. M., Domanico, P. L. & Stray, S. J. (2007). *In vitro* screening for molecules that affect virus capsid assembly (and other protein association reactions). *Nature Protoc.* **2**, 490–498.
80. Bernd, J. (2004). *Practical Handbook on Image Processing for Scientific and Technical Applications*, 2nd edit. CRC Press, Boca Raton, FL.
81. Duus, J. O., Meldal, M. & Winkler, J. R. (1998). Fluorescence energy-transfer probes of conformation in peptides: The 2-aminobenzamide/nitrotyrosine pair. *J. Phys. Chem. ser. B*, **102**, 6413–6418.
82. Park, C., Kelemen, B. R., Klink, T. A., Sweeney, R. Y., Behlke, M. A., Eubanks, S. R. & Raines, R. T. (2001). Fast, facile, hypersensitive assays for ribonucleolytic activity. *Methods Enzymol.* **341**, 81–94.

J Biol Inorg Chem (2008) 13:371–383
DOI 10.1007/s00775-007-0325-8

ORIGINAL PAPER

Systematic development of computational models for the catalytic site in galactose oxidase: impact of outer-sphere residues on the geometric and electronic structures

Dalia Rokhsana · David M. Dooley · Robert K. Szilagyi

Received: 24 October 2007 / Accepted: 14 November 2007 / Published online: 4 December 2007
© SBIC 2007

Abstract A systematic *in silico* approach has been employed to generate sound, experimentally validated active-site models for galactose oxidase (GO) using a hybrid density functional, B(38HF)P86. GO displays three distinct oxidation states: oxidized [Cu(II)–Y•]; semireduced [Cu(II)–Y]; and reduced [Cu(I)–Y]. Only the [Cu(II)–Y•] and the [Cu(I)–Y] states are assumed to be involved in the catalytic cycle, but their structures have not yet been determined. We have developed several models (**1**–**7**) for the [Cu(II)–Y•] state that were evaluated by comparison of our computational results with experimental data. An extended model system (**6**) that includes solvent molecules and second coordination sphere residues (R330, Y405, and W290) is essential to obtain an experimentally correct electronic structure of the active site. The optimized structure of **6** resulted in a five-coordinate Cu site with a protein radical centered on the Tyr–Cys cofactor. We further validated our converged model with the largest model (**7**) that included additional outer-sphere residues (Q406, H334, Y329, G513, and T580) and water molecules. Adding these residues did not affect significantly the active site's electronic and geometric structures. Using both **6** and **7**, we explored the redox dependence of the active-site structure. We obtained four- and three-coordinate Cu sites for [Cu(II)–Y] and [Cu(I)–Y] states, respectively, that corroborate well with the experimental data. The relative energies of these states were validated by a comparison with experimental redox potentials. Collectively, our computational GO models well reproduce the physicochemical characteristics of the individual states, including their redox behaviors.

Keywords Galactose oxidase · Copper/protein radical active site · Computational biomimetic model · Density functional theory · Spectroscopic calibration of theory

Introduction

Galactose oxidase (GO) is an extracellular, monomeric copper protein of molecular mass approximately 68 kDa, isolated from the fungus *Dactylium dendroides* [1]. It is composed of three distinct domains, organized in an exclusively β-sheet structure with short turns. One domain contains a two-electron redox unit with a mononuclear copper ion and a cross-linked Tyr–Cys radical cofactor that are responsible for its catalytic activity. GO catalyzes the two-electron oxidation of a broad range of primary alcohols (e.g., D-galactose, D-galactoseamine, dihydroxyacetone) to the corresponding aldehydes, concomitantly reducing dioxygen to hydrogen peroxide [1]. GO displays three distinct oxidation states: oxidized ([Cu(II)–Y•], green); semireduced ([Cu(II)–Y], blue); and reduced ([Cu(I)–Y], colorless). The oxidized and the reduced states are catalytically important in the two-electron redox reaction during the catalytic cycle [2].

Crystal structures are available for the [Cu(II)–Y] state at pH 4.5 and 7.0 (1.7- and 1.9-Å resolution, respectively).

Electronic supplementary material The online version of this article (doi:10.1007/s00775-007-0325-8) contains supplementary material, which is available to authorized users.

D. Rokhsana · D. M. Dooley (✉) · R. K. Szilagyi (✉)
Department of Chemistry and Biochemistry,
Montana State University,
Bozeman, MT 59717, USA
e-mail: dmdooley@montana.edu

R. K. Szilagyi
e-mail: szilagyi@montana.edu

Both structures suggest a distorted square-pyramidal Cu(II) coordination geometry with an axial Y495 ligand and four equatorial ligands of H496, H581, H_2O (pH 7.0), or acetate (pH 4.5), and the unique cross-linked Tyr–Cys cofactor (Fig. 1) [3]. Detailed, atomic-resolution structures of both catalytically important states ([Cu(II)-Y•] and [Cu(I)–Y]) are not yet available owing to their inherent instability [4]; however, structural characterization of these redox states is essential to understand the molecular mechanism of GO.

Spectroscopic studies have provided considerable insights into the copper coordination environment in GO [2, 5]. On the basis of extended X-ray absorption fine structure (EXAFS) studies, it was suggested that the coordination environment of the oxidized [Cu(II)–Y•] is similar to that of semireduced [Cu(II)–Y] state (Fig. 1) [5]. A three-coordinate Cu site is suggested for the reduced [Cu(I)–Y] state with an average of two Cu–N bond distances and a Cu–O bond distance of 1.95 and 1.97 Å, respectively [6]. In contrast to the geometric structures, the electronic absorption spectra [7] of the GO redox states differ significantly. The [Cu(II)–Y] state shows weak absorption bands at 450 and 620 nm, whereas the [Cu(II)–Y•] state displays multiple, intense electronic transitions that span the visible/near-infrared regions with maximum intensity at 445 and 800 nm. These bands are assigned as ligand-to-metal charge transfer (LMCT) and ligand-to-ligand charge transfer (LLCT) transitions, respectively [4, 8]. The fully reduced [Cu(I)–Y] state shows no characteristic signatures in the electronic spectral region owing to the filled *d* orbitals of the Cu(I) ion and the nonradical form of the Tyr–Cys cofactor.

A computational study on the catalytic cycle of GO has been previously reported employing a semiquantitative analysis of the experimental data and relevant model systems of GO [9]. More recently, quantum mechanical/molecular mechanics Car–Parrinello simulations of GO and its synthetic active-site analogs were carried out [10] using a small computational model and hybrid density functional theory (DFT; B3LYP) [11]. These results provided new perspectives for developing structure–reactivity correlations for GO. This work suggested a distorted tetragonal Cu site for the [Cu(II)–Y•] state with longer Cu–O (H_2O and Y495) distances of 2.6–3 Å, than observed in the EXAFS (1.96 Å) [5, 10]. Notably, these calculations indicated that the protein radical was located on the axial ligand Y495 rather than the equatorial Tyr–Cys cofactor [11]. However, electron paramagnetic resonance studies have demonstrated that the protein radical is on the cross-linked Tyr–Cys cofactor for [Cu(II)–Y•] [12–15]. In addition, the magnetic susceptibility study revealed that the protein radical is antiferromagnetically coupled with the Cu(II) ion, resulting in a diamagnetic ground state for [Cu(II)–Y•] [2]. Therefore, further investigations are needed to reconcile the experimental findings with the calculated electronic and geometric structures of the active site. Moreover, there is a growing interest in designing synthetic biomimetic analogs [16–22] based on the GO active-site architecture for stereoselective alcohol oxidation that can be aided by gaining more insights into the structures of the important oxidation states of GO, and the factors that govern its reactivity.

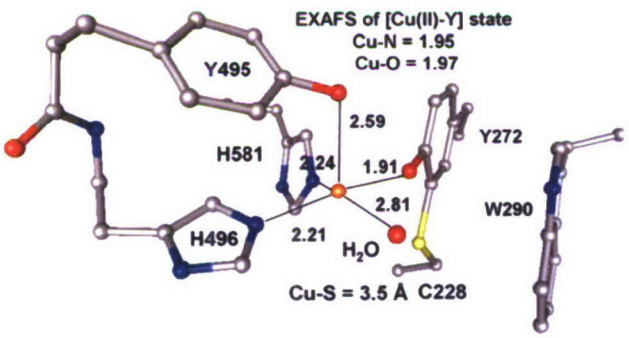


Fig. 1 Crystal structure of the active site of galactose oxidase (GO) at 1.9 Å resolution in the [Cu(II)–Y] state including residues known to affect catalytic activity with selected Cu–ligand distances (Protein Data Bank ID 1GOG) [3]. The average Cu–N and Cu–O extended X-ray absorption fine structure (*EXAFS*) distances for oxidized [Cu(II)-Y•] are 1.97 and 1.96 Å, respectively [5]

We describe here a systematic *in silico* approach for developing accurate active-site models of GO to study the coordination geometries and electronic structures of all three distinct redox states. A spectroscopically calibrated hybrid density functional [B(38HF)P86] was employed, which has been validated for mononuclear copper proteins [23, 24]. Our model-building strategy was designed to evaluate the importance of each inner-shell and outer-shell residue on the geometric and electronic structures of the active site. Each model (Fig. 2) of the oxidized GO was rigorously evaluated by comparing the calculated and experimental structures, spin-density distribution, singlet–triplet energy gap, and electronic absorption spectra. The calculated [Cu(II)–Y] and [Cu(I)–Y] states were obtained by one-electron reduction of the optimized [Cu(II)–Y•] and [Cu(II)–Y] states, respectively. In addition to using the molecular orbital coefficients of [Cu(II)–Y•], [Cu(II)–Y] and [Cu(I)–Y] models as initial orbital guesses for [Cu(II)–Y] (blue solid), [Cu(II)–Y•] (green dotted arrow) and [Cu(I)–Y] (black solid arrow), [Cu(II)–Y] (blue dotted arrow), respectively, we carried out optimizations starting from the crystal structure of the [Cu(II)–Y] form with the ionic fragments [25] (Fig. 3). These fragments correspond to the valence-bond description of amino acid ligands and the Cu(II) ion. These procedures were performed to evaluate the robustness of the optimization method and to investigate protein-strain effects.

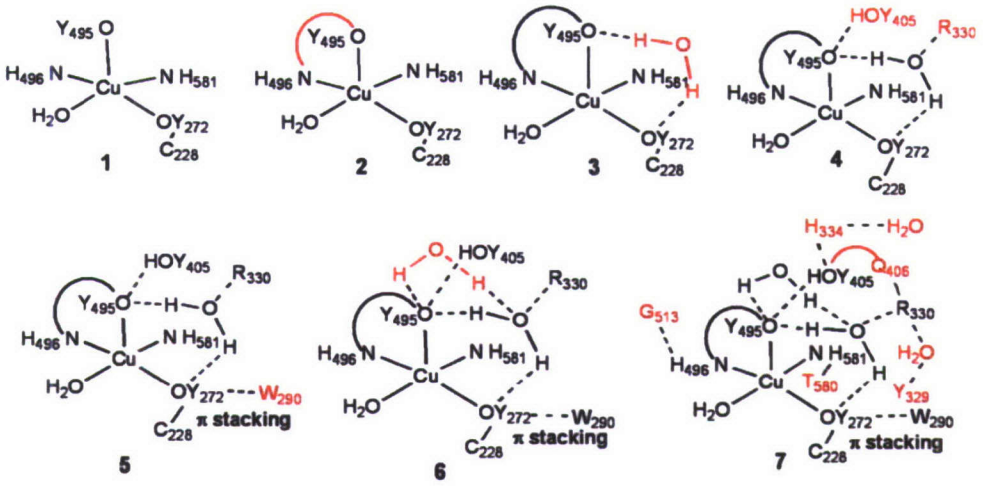


Fig. 2 Computational models of GO. Changes between two consecutive models are shown in *red*

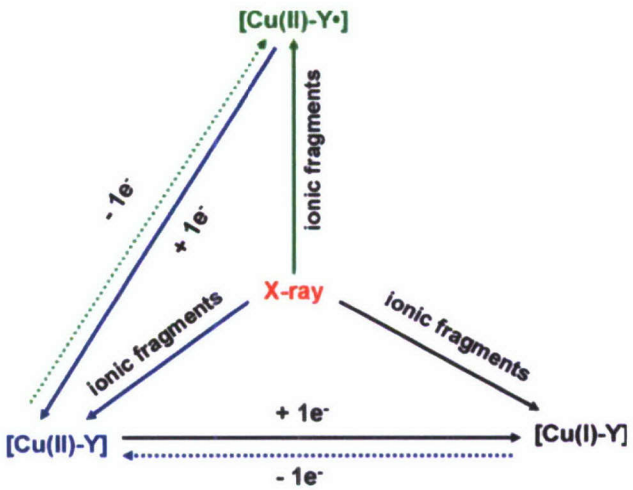


Fig. 3 Systematic mapping of the potential energy surface of redox states for the GO active site

Overall, our study quantitatively shows that the protein environment has significant influences on the properties of the active site, consistent with previous site-directed mutagenesis studies of second coordination sphere residues (W290H, W290F, W290G, and Y495F) [26, 27]. It is interesting to note that our computational models were developed without following the site-directed mutagenesis study, rather allowing for the model to reveal the residues that are essential in determining the active-site structure. We have already communicated the results of a converged model for the structure of oxidized [Cu(II)–Y•] GO [28]. In this paper, we provide the full details of our model-building strategy for oxidized [Cu(II)–Y•] GO, the semireduced [Cu(II)–Y], and the reduced [Cu(I)–Y] states (Fig. 3). We further validated these models by comparing the redox potentials for [Cu(II)–Y•]/[Cu(II)–Y] and [Cu(II)–Y]/[Cu(I)–Y] with the experimental values [29]. Our work lays a strong foundation for investigating the molecular mechanism of GO. It also provides a demonstrative example for using a systematic *in silico* model-building approach for constructing structurally and functionally accurate computational models of metalloprotein active sites.

Computational details

A spectroscopically calibrated, hybrid density functional [B(38HF)P86] [24] was employed in our computational study using the Gaussian03 package [30] on a cluster of Intel Xeon EM64T servers. The Becke88 exchange and Perdew86 correlation functional were used with 38% of the total density functional exchange replaced with Hartree–Fock exchange, giving an accurate bonding description for mononuclear copper proteins [23]. The triple-ζ (VTZ*) [31] and double-ζ with polarization [6-31G(d)] [32–34] Gaussian-type all-electron basis sets were employed in all calculations, which form a converged basis set for Cu(II)-containing systems [24, 35].

The generalized ionic fragment approach [25] was employed for model building. First, the molecular orbitals of each residue as ionic ligand fragments were calculated in their crystallographic positions in the [Cu(II)–Y] form (Protein Data Bank ID 1GOG) [3]. Second, these molecular orbitals were combined to give a specific spin-polarized wave function for the complete active-site models. Third, geometry optimizations were performed using spin-unrestricted B(38HF)P86 to maintain the open-shell singlet character of the wave function. As summarized in Fig. 2, our computational models span from a truncated model (H496 and H581 modeled as imidazole, Y495 modeled as phenol, Y272 modeled as thiophenol and H_2O) to an extended model system with 214 atoms, which includes the α-carbon of each residue coordinated to the Cu(II); the

protein backbone between Y495 and H496; an additional H_2O (HOH294) within hydrogen-bonding distance from both Tyr residues (approximately 3 Å in the crystal structure); distal residues Y405, Q406, R330, W290, G513, T580, H334, and Y329, and additional crystal water molecules.

The His residues were modeled as charge-neutral, Y495 as tyrosinate (negatively charged), Y272 as neutral tyrosyl radical, Y405 as protonated Tyr (neutral), R330 as protonated guanidium (positively charged), and all other residues as charge-neutral unless otherwise noted. The total charges of models **1**–**3** were +1, 0, and −1 and those of **4**–**7** were +2, +1, and 0 for [Cu(II)–Y•], [Cu(II)–Y], and [Cu(I)–Y], respectively. For models **2**–**7**, α-carbons of each residue were fixed at their crystallographically determined positions during optimization unless otherwise noted. Geometry optimizations were performed using the redundant coordinate system and Berny optimization algorithm [36]. The convergence criteria were set to 0.002 hartree/(au or rad) and 0.01 au or radian RMS change in energy gradients and internal coordinate displacements, respectively. Owing to the relatively lax convergence criteria, geometry optimizations were terminated only when stationary points were firmly located in consecutive calculations. For larger models, this procedure avoided trapping in a higher-energy local minimum. Atomic spin densities were derived from Mulliken population analysis.

The relative energies of Kohn–Sham orbitals from DFT calculation have been shown to provide reasonable estimates of absorption bands for copper proteins [37]. Time-dependent DFT (TD-DFT) [38] calculations were also performed to obtain an improved description of electronic transitions for these models. Antiferromagnetic coupling magnitudes were determined from the energy difference between the optimized singlet and triplet states for each model using spin Hamiltonian $J \cdot S_A S_B$ ($E_T - E_S = 2$ $J \cdot S_{Cu(II)} \cdot S_{Tyr\bullet-Cys}$). Spin projection of the broken-symmetry DFT solution to obtain E_S was done by the method suggested by Davidson and Clark [39].

Redox potentials were calculated from the sum of the ionization potentials and solvent reorganization energies relative to the normal hydrogen electrode at −4.43 eV. The solvent reorganization energies were implicitly considered by polarizable continuum calculations using various dielectric constants and solvent radii of 1.385 Å (water). For the proton source in the protonated structures, we used the hydronium ion surrounded by four explicit water molecules. This was further embedded into a standard water-based polarizable continuum. In order to obtain accurate energy for the proton source in the latter system, we used a triple-ζ-quality basis set with both polarization and diffuse functions (6-311+G*).

Results and discussion

First, we demonstrate the importance of using a systematic model-building procedure on the oxidized form of GO, in order to capture all the significant first and second coordination sphere effects. We selected the oxidized state of GO owing to the availability of numerous spectroscopic measurements that are utilized to validate our computational models. Starting from the simplest active-site model (**1**), we systematically extend this to the converged structural model (**6**) that reproduces most of the experimental data. We show that a considerably larger model (**7**) does not improve the calculated electronic and geometric structural data in comparison with experiment relative to model **6**. Using models **6** and **7**, we describe the structural changes that accompany the relevant redox reactions, and for **6** we also explicitly evaluate the redox potentials.

Systematic development of a converged structural model for oxidized GO

Model 1

The smallest GO model (**1**) in our computational studies comprises four equatorial ligands, H496, H581, the cross-linked Tyr–Cys cofactor, and water, and an axial ligand, Y495. Both His residues are modeled as imidazoles (charge-neutral), the axial Y495 as phenolate (negatively charged), and the cross-linked thioether–substituted Tyr as a CH_3S–substituted phenyl radical. This model is similar to that used in an earlier computational study [11]. The optimized structure of oxidized **1** with selected Cu–ligand distances is shown in Fig. 4a. Model **1** can be characterized as a slightly distorted octahedral Cu site with a phenol ligand in the axial position at 2.29 Å distance. The sulfur atom of the equatorial CH_3S–substituted phenyl is located axially, *trans* to the phenol ligand at a distance of 2.97 Å. Two imidazoles, the substituted phenolate, and the H_2O are forming the equatorial plane with Cu–ligand distances ranging from 1.91 to 2.10 Å, and with *cis* and *trans* ligand–Cu–ligand angles of 79–100° and 165 and 179°, respectively (Fig. 4a). The aromatic ring of the axial phenol is rotated more than 90° from it is crystallographic position, which suggests the importance of steric constraints from the protein backbone in positioning the aromatic ring almost perpendicular to the equatorial plane (Fig. 1).

The electronic structure of oxidized **1** (Fig. 4b) reveals 0.74 e^- spin density located on the Cu ion, and the rest of the α-spin is delocalized onto the equatorial ligands: both His ligands (0.06 e^- on each εN), phenolate (0.09 e^-),

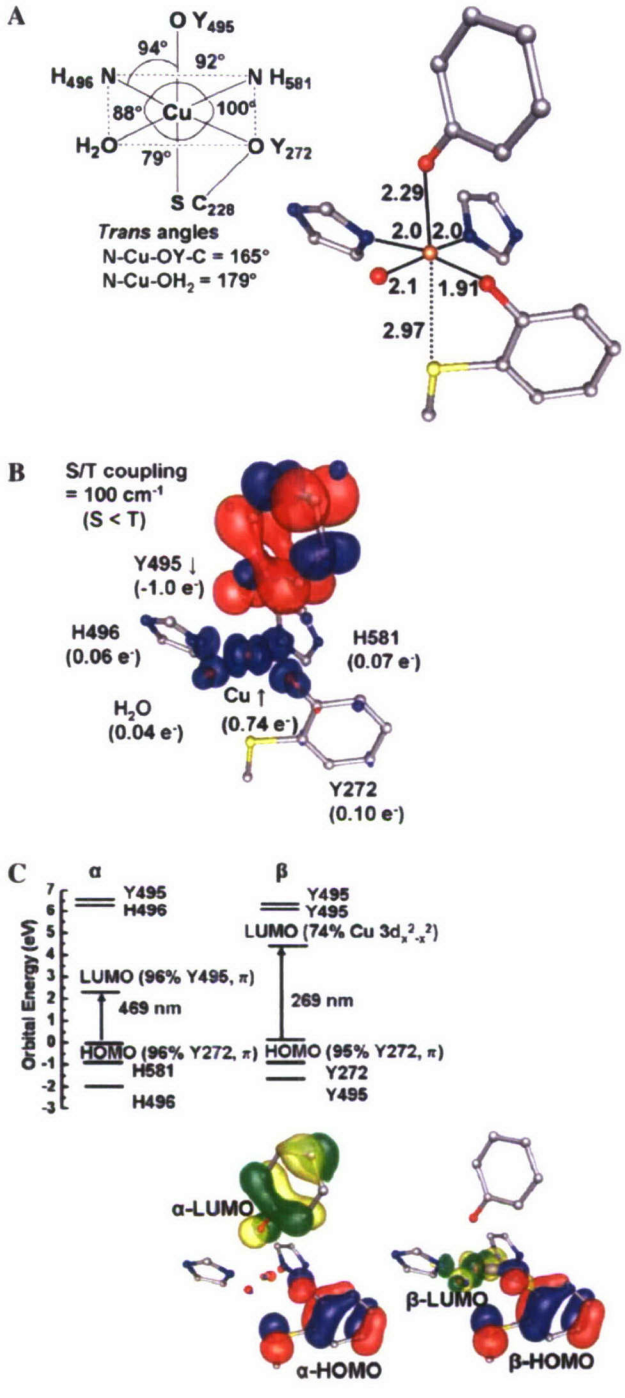


Fig. 4 Optimized structure of model **1**. **a** Cu–ligand distances (angstroms) and ligand–Cu–ligand angles (degrees). **b** Spin-density plot (isosurface of ±0.003) and fragment spin densities, and **c** Kohn–Sham molecular orbital energies (*top*) and the highest occupied molecular orbital (*HOMO*) and the lowest unoccupied molecular orbital (*LUMO*) (isosurface of ±0.05) of both α and β manifolds (*bottom*)

and H_2O (0.04 e^-). The β-spin is mainly confined to the axial phenol ring (Fig. 4b), which is inconsistent with experimental findings [12–15]. Since our result agrees with previous computational studies obtained employing a different density functional [10, 11], the location of the spin in this minimal model is likely determined by the computational model and not by the selected functional or basis set.

We have extensively analyzed various initial triplet and singlet state structures of **1** by using a generalized ionic fragment approach [25] with unpaired electrons located at different fragments, such as Cu(II)/substituted phenolate, Cu(II)/axial phenol, and substituted phenolate/axial phenol with a Cu(I) center in the last case (see "Computational details"). In all cases, the structures converged to the same spin-density distribution as described above. Furthermore, the calculated antiferromagnetic coupling ($J \sim 100\ cm^{-1}$) is considerably smaller than the experimental antiferromagnetic coupling ($J > 200\ cm^{-1}$). In addition, the calculated electronic transitions from a Kohn–Sham orbital analysis are as follows: 266 nm, substituted phenolate to Cu(II) $3d_{x^2-y^2}$ ($\pi \rightarrow \sigma^*$), and 469 nm, substituted phenolate to phenyl radical ($\pi \rightarrow \pi$) (Fig. 4c), which are also in disagreement with the experimental electronic transitions observed for [Cu(II)–Y•] GO. These results obtained for the minimal coordination chemistry model of the active site suggest that the protein backbone between Y495 and H496 and interactions among amino acid side chains are important in constraining the positions of Cu(II) ligands.

Model 2

Model **1** was extended by adding the backbone linkage between Y495 and H496 and the α-carbons of each residue to take into account steric constraints from the protein backbone. Optimization of **2** was performed by freezing all α-carbons at their crystallographic positions. This resulted in a tetragonal copper center similar to in **1**, but the axial Y495 and C228 ligands moved away considerably from the copper center (Cu–O(Y495) distance 3.47 Å, and Cu–S(C228) distance 3.56 Å; Fig. 5a). Also, the position of the equatorial H_2O is slightly above the equatorial plane and the H_2O forms a hydrogen-bonding interaction with the oxygen atom of the axial Y495 (O–O distance 2.67 Å). The average bond distances of 1.98 Å (Cu–N distance 2.00 Å and Cu–O distance 1.94 Å) in **2** are in reasonable agreement with the distances obtained from EXAFS study (Cu–N distance 1.97 Å and Cu–O distance 1.96 Å).

The calculated electronic structure of **2** suggests a slightly more covalent bonding than in **1** involving a $3d_{x^2-y^2}$ orbital with 0.67 e^- α-spin density located on the Cu(II) ion, and the remaining 0.34 e^- spin density delocalized onto the σ-bonded equatorial ligands [Y272 (0.15 e^-), H496 (0.06 e^-), H581 (0.08 e^-), and H_2O (0.05 e^-)]. The protein radical (β-spin) is mainly localized

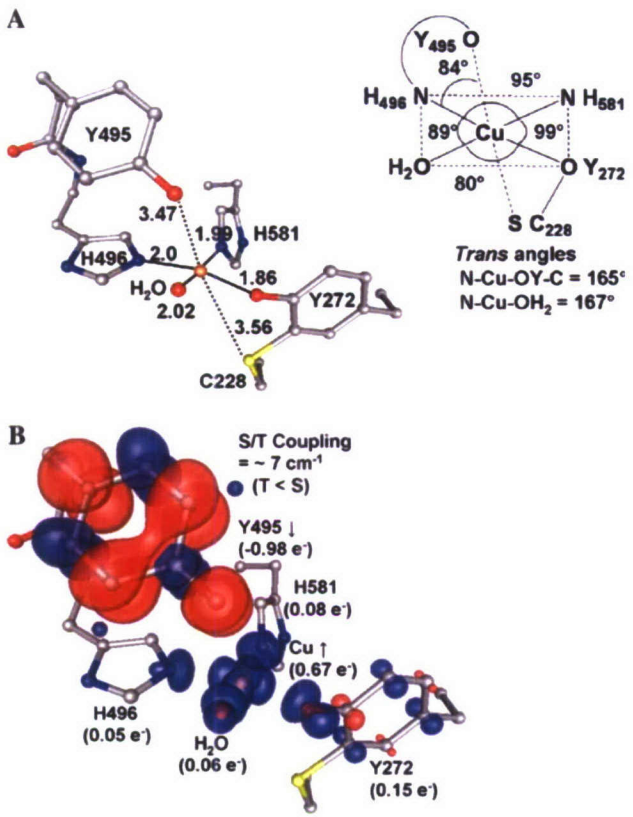


Fig. 5 Optimized structure of model **2**. **a** Cu–ligand distances (angstroms) and ligand–Cu–ligand bond angles (degrees). **b** Spin-density plot (isosurface of ±0.003) and fragment spin densities

onto a π orbital of axial Y495 (Fig. 5b). The calculated antiferromagnetic coupling J (approximately 7 cm^{-1}) is almost negligible, yielding nearly degenerate triplet and singlet states. Hence, the calculated electronic structure of **2** is inconsistent with the experimental data in the location of protein radical, the negligible singlet–triplet energy gap, and the inaccurate electronic transitions (Fig. S1).

Model 3

We reasoned that for changing the location of the Tyr radical from axial to equatorial it would be desirable to reduce the nucleophilicity of the axial Tyr (Y495) ligand and thus destabilize a radical at this position. Accordingly, model **2** was extended via addition of a hydrogen-bonding water molecule, which is resolved in the crystal structure [3]. The optimized **3** with the constraints as for **2** resulted in a tetragonal pyramidal copper site with all equatorial ligands located within 1.89–2.0 Å distances; the shortest being the Cu–O(Y272) (Fig. 6a). The axial Y495 is located at 2.49 Å from the copper center, and the Tyr ring is almost parallel to the equatorial plane. The electronic structure of **3** shows small changes in the electron spin distribution

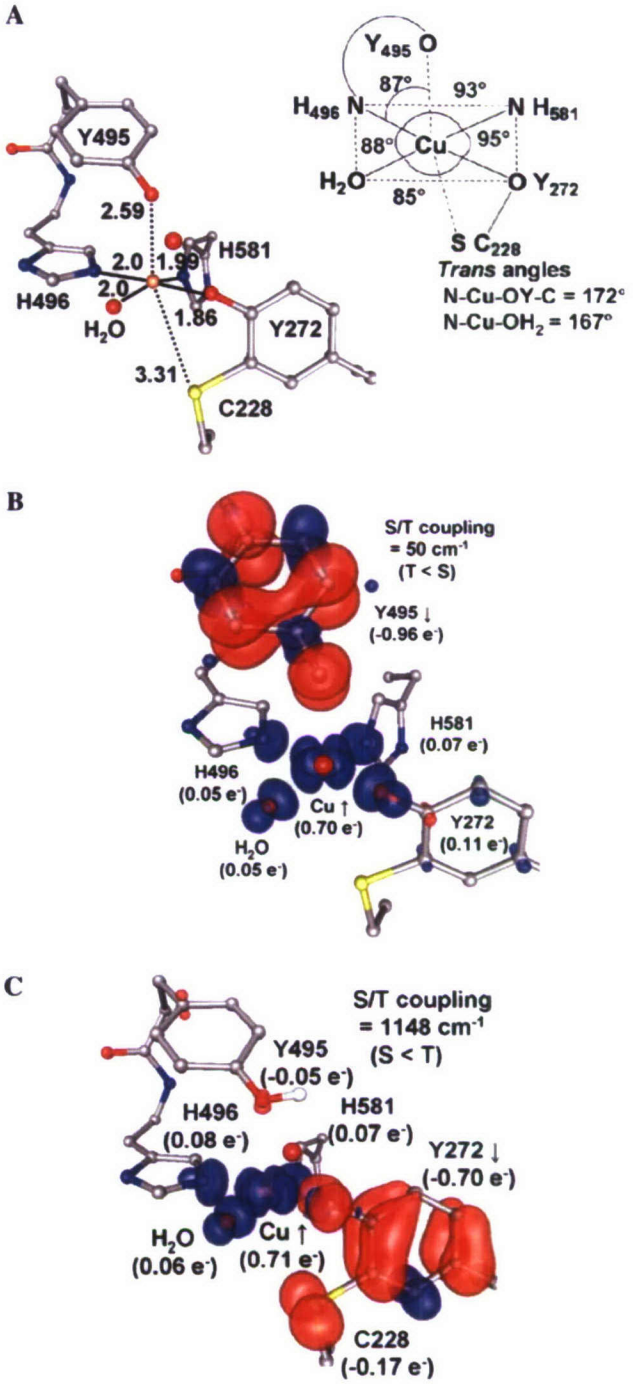


Fig. 6 Optimized structure of model **3**. **a** Cu–ligand distances (angstroms) and ligand–Cu–ligand angles (degrees). **b** Spin-density plot (isosurface of ±0.003) and **c** spin-density plot of Y495 protonated in [Cu(II)–Y•] model **3** (isosurface of ±0.003)

despite the geometric difference relative to **2**, and results in a triplet ground state (Fig. 6b).

In order to further probe the idea that the reduced nucleophilicity of axial Y495 could tune the location of the

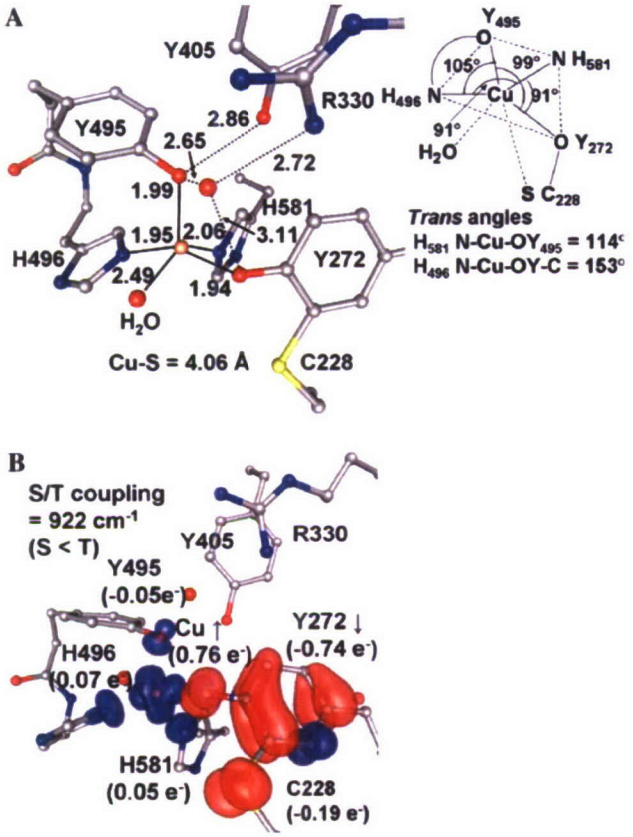


Fig. 7 Optimized structure of model **4a**. **a** Cu–ligand distances (angstroms, *solid lines*) and hydrogen-bonding interaction (angstroms, *dotted lines*) and ligand–Cu–ligand angles (degrees). **b** Spin-density plot (isosurface of ±0.003)

protein radical, we evaluated the possibility of the O(Y495) being protonated. It is worth noting that Himo et al. [11] also reported that in the substrate-bound form of oxidized GO the radical is transferred to the equatorial Tyr–Cys cross-link simultaneously with the proton transfer to O(Y495) from the substrate. In our case without substrate present, a single point energy calculation for the protonated **3** resulted in a spin distribution that is consistent with experiment. For the first time, the radical (unpaired β-spin electron) is located on an out-of-plane π orbital of the Tyr–Cys cofactor and the unpaired α-spin electron is localized on the Cu $3d_{x^2-y^2}$ orbital and the in-plane orbitals of the covalently bound equatorial ligands (Fig. 6c). This result motivated us to assess the possibilities of other hydrogen-bonding residues around the oxygen atom of Y495, leading to model **4**.

Model 4

The oxygen atom of the Tyr (Y405) residue is located at approximately 3.25 Å away from the O(Y495) and it is well positioned for hydrogen bonding with Y495. Furthermore, a positively charged Arg residue (R330) is located within hydrogen-bonding range of H_2O (294), which could stabilize the position of this water molecule. Note that the position of H_2O (294) in model **3** changed considerably by moving toward Y272 ($d_{O-O(Y272)}$ = 2.66 Å) and forming a stronger hydrogen-bonding interaction with Cu-coordinated H_2O (d_{O-O} = 2.58 Å) relative to the crystal structure ($d_{O-O(Y272)}$ = 3.02 Å and d_{O-O} = 3.01 Å). Therefore, model **4** consists of protonated second coordination sphere residues Y405 and R330 in addition to the residues included in **3** (Fig. 2). Initially, geometric optimization of **4** was performed with fixed crystallographic positions for α-carbons of each residue (**4a**). As shown in Fig. 7a, model **4a** has a tetragonally distorted pyramidal coordination geometry with H_2O in an axial position at 2.49 Å, whereas Y495 now occupies an equatorial position. The equatorial plane is defined by the four shortest Cu–ligand distances (1.94–2.06 Å), with *cis* and *trans* ligand–Cu–ligand bond angles of 91–105° and 114 and 153°, respectively (Fig. 7a). With respect to the electronic structure, this means that the equatorial plane containing the $3d_{x^2-y^2}$ orbital is rotated by 90° relative to that in models **1**–**3**.

The calculated electronic structure of the oxidized [Cu(II)–Y•] state of **4** suggests a metal-centered bonding description involving a $3d_{x^2-y^2}$ orbital with 0.76 e^- α-spin density located on the Cu(II) ion, and the remaining 0.24 e^- is delocalized evenly onto the σ-bonded equatorial ligands (Fig. 7b). The protein radical (Tyr•–Cys) has approximately 0.94 e^- β-spin density in an out-of-plane π orbital, including approximately 0.20 e^- spin delocalization onto the sulfur atom of C228. The possible involvement of S orbitals in the protein radical predicted by this model parallels the conclusions from studies of biomimetic GO models [14, 40].

An alternative optimization of oxidized **4** was performed with additional distance constraints from EXAFS [both Cu–N distance 1.97 Å and Cu–O(Y272) distance 1.96 Å] in addition to fixing α-carbons at their crystallographic positions (**4b**). Interestingly, this approach produced a trigonally distorted tetrahedral Cu(II) geometry with ligand–Cu–ligand angles spanning from 99 to 125° (Fig. 8a), which is quite different from the situation for **4a**, and resembles the angular parameters of blue copper proteins [37]. The equatorial H_2O in **4b** is quite distant from the copper center at 3.42 Å. Both models (**4a** and **4b**) agree with EXAFS results but differ in their electronic structures. In **4b** the Cu(II) spin density was increased to approximately 0.86 e^- α-spin, corresponding to less covalent bonding (Fig. 8b). The delocalization of the protein radical (Tyr•–Cys) onto the sulfur atom of C228 was also reduced by about 4% (Fig. 8b). The calculated antiferromagnetic coupling constants are 923 and 39 cm^{-1} for **4a** and **4b**,

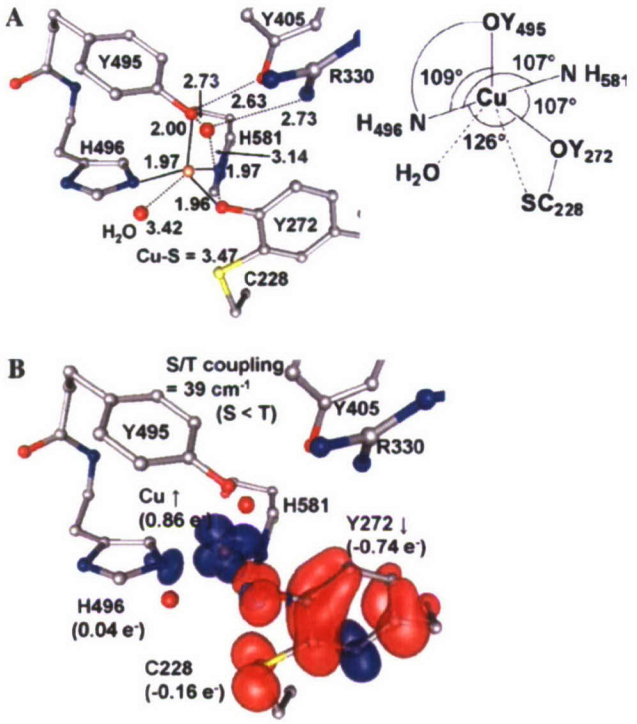


Fig. 8 Optimized structure of model **4b**. **a** Cu–ligand distances (angstroms, *solid lines*), hydrogen-bonding interactions (angstroms, *dotted lines*), and ligand–Cu–ligand angles (degrees). **b** Spin-density plot (isosurface of ±0.003)

respectively. It is also important to note that model **4a** is approximately 68 kJ mol^{-1} lower in energy than **4b**; therefore, model **4a** was considered a more reasonable model than **4b**. TD-DFT analysis was performed for **4a** and predicted three intense features at 766, 390, and 372 nm, giving rather close agreement with the experimental optical transition. They can be assigned to tyrosinate ($\pi_3 \rightarrow \pi_4$), Y272 ($\pi_2 \rightarrow \pi_4$), and Y272 (π_3) $\rightarrow \sigma^*$ of Cu(II) $3d_{x^2-y^2}$ transitions, respectively.

Model 5

While **4a** can already be considered as a reasonable computational model for the [Cu(II)–Y•] state of the GO active site, it lacks a key residue (W290) that has been shown to influence the catalytic activity by affecting the generation of the tyrosyl radical (Y272) in GO [41]. In the crystal structure, it also appears that the conformation of Y272 is affected by a π stacking interaction with W290. This stacking interaction was included in model **5** to evaluate the effect of Trp on the geometric and the electronic structure of the [Cu(II)–Y•] state. The optimized structure of oxidized **5** resulted in a square-planar copper site with all equatorial ligands located within 1.87–2.04 Å distance, the shortest being Cu–O (Y272). The axial Y495 is located 3.17 Å from the copper center, and almost parallel to the equatorial plane of the copper center. Unexpectedly, R330 and Y405 residues were distorted significantly in the optimized structure relative to the crystal structure, in contrast to the results for **4**. W290 was also dislocated from the stacking plane. Moreover, the electronic structure of **5** is inconsistent with experiment since the protein radical was mainly localized on a π orbital of the axial Y495 (Fig. S2).

These results for model **5** indicate that the π–π stacking interaction of the cross-linked Tyr–Cys cofactor and W290 can have a significant influence on the electronic structure of the [Cu(II)-Y•] state. We inferred that the protein environment must modulate the interaction between W290 and the Tyr–Cys cofactor. Inspection of the crystal structure revealed that both the axial Y495 and the indole nitrogen of W290 are partially exposed to solvent. Because our computational work already revealed that hydrogen bonding can modulate the electron spin distribution in the GO active site, we introduced an explicit water solvent molecule that is within hydrogen-bonding distance of the indole nitrogen of W290 or of the oxygen of Y495 in separate models. Only the latter hydrogen-bonding interaction resulted in a protein radical on the cross-linked Y272–C228 cofactor defining our model **6**.

Model 6

The spin-polarized wave function of **6** was obtained from two different starting points owing to issues in self-consistent-field convergence: merging ionic wave functions from individual fragments; or adding the W290 residue and the water molecule to a covalent, spin-polarized wave function of **4**. The latter resulted in a lower-energy wave function (atomic spin densities of 0.74 e^- on Cu, −0.60 e^- on Y272, −0.15 e^-on Y495, and −0.10e^- on W290) by approximately 104 kJ mol^{-1} relative to the former wave function (atomic spin densities of 0.99 e^- on Cu, −0.75 e^- on Y272, −0.14 e^- on Y495, and −0.09e^- on W290). It is worth noting that these results demonstrate well potential pitfalls for finding the ground-state electronic state in large computational models from single-point calculations. The generalized ionic fragment approach may avoid these pitfalls by allowing for generating all possible valence-bond structures as initial wave functions in a systematic and reproducible manner. Optimization of **6** starting from the lower-energy spin-polarized wave function was performed with additional constraints of R330 (δN, εC, and χN) in order to avoid the unreasonable distortion of R330 that was observed in **5**. The optimized structure of **6** reveals a five-coordinated square-pyramidal Cu coordination geometry with Y495 in an axial position and the other ligands in

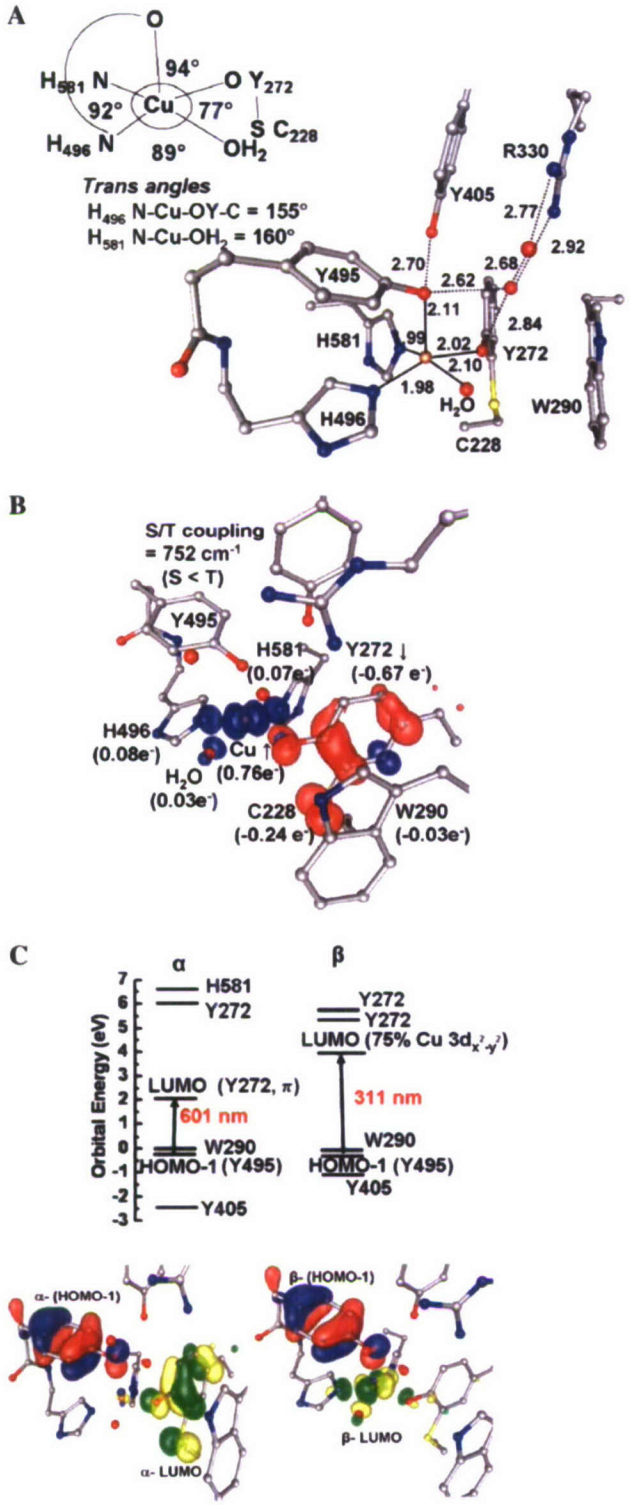


Fig. 9 **a** Optimized structure of model **6** with Cu–ligand (angstroms, *solid lines*), hydrogen-bonding (angstroms, *dotted lines*) distances and ligand–Cu–ligand angles (degrees). **b** Spin-density plot (isosurface of ±0.003). **c** Relative energies of Kohn–Sham molecular orbitals (*top*) and HOMO-1 and LUMO (isosurface of ±0.05) of both α and β manifolds (*bottom*)

equatorial positions with *cis* and *trans* ligand–Cu–ligand bond angles of 77–94° and 155 and 159°, respectively (Fig. 9a). The overall coordination environment of the copper site in [Cu(II)–Y•] is comparable to that in the crystal structure of the [Cu(II)–Y] state with some deviations in the Cu–ligand bond distances. It is interesting to note that the average Cu–N and Cu–O distances (1.99 and 2.06 Å) also slightly deviate from the corresponding EXAFS values. The calculated electronic structure of the oxidized [Cu(II)-Y•] state of **6** suggests a metal-based bonding description involving a $3d_{x^2-y^2}$ orbital with 0.76 e^- α-spin density located on the Cu(II) ion. The Tyr•–Cys radical has approximately 0.91e^- β-spin density in an out-of-plane π_4 orbital (Fig. 9b; for orbital labeling see Fig. S3). This includes approximately 0.24 e^- spin delocalization onto the sulfur atom of C228. The π–π stacking between W290 and Tyr•–Cys (approximately 3.7 Å), and constraints from α-carbon atoms of Y272, C228, and W290 maintain a cofactor orientation (Cu–O–C bond angle, $\alpha = 125°$ and dihedral angle between equatorial plane of Cu coordinated ligands and tyrosyl ring plane, $\beta \sim 82°$) that results in a diamagnetic ground state [22].

It is important to mention that the proton on the indole N of W290 is bent and about 0.03 e^- spin density is located on the Trp ring. A Kohn–Sham molecular orbital energy diagram for the oxidized model is shown in Fig. 9c. The electronic transitions between HOMO-1 and LUMO orbitals at 311 and 601 nm in the spin-down (β) and the spin-up (α) manifolds, respectively, can be assigned to LMCT and LLCT excitations. The former corresponds to the tyrosinate (Y495) to Cu $3d_{x^2-y^2}$ orbital,

$\pi_4 \rightarrow \sigma^*$, and the latter to tyrosinate (Y495) to tyrosyl (Y272) radical, $\pi_4 \rightarrow \pi_4$ (Fig. 9c). TD-DFT analysis gave an improved agreement between the experimental and calculated band positions and intensities. Nonzero oscillator strengths (f) were obtained for transitions at 795 nm ($f = 0.02$, radical Y272 $\pi_3 \rightarrow \pi_4$), 400 nm ($f = 0.04$, radical Y272 $\pi_2 \rightarrow \pi_4$), and 359 nm [$f = 0.05$, H496 (π) and Y272 (π_4) to σ^* of Cu(II) $3d_{x^2-y^2}$]. These are consistent with assignments based on electronic absorption and resonance Raman measurements [42, 43] except that the TD-DFT calculations do not predict substantial contributions from Y495, as inferred from the resonance Raman experiments. The calculated antiferromagnetic coupling is about 752 cm^{-1} for **6**.

Model 7

In order to demonstrate that model **6** is indeed a converged structural model for oxidized GO, which already provides

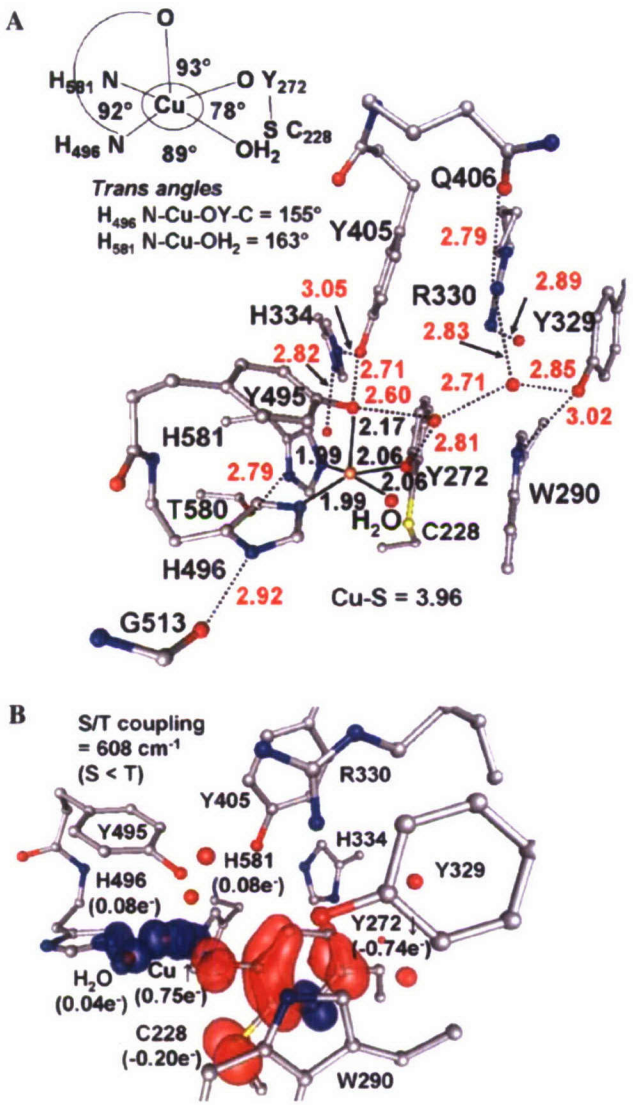


Fig. 10 **a** Optimized structure of [Cu(II)–Y•] model **7** with Cu–ligand (angstroms, *solid lines*) and hydrogen-bonding (angstroms, *dotted lines*) distances and ligand–Cu–ligand angles (degrees). **b** Spin-density plot (isosurface of ±0.003)

good agreement with experiment, we further extended this computational model to 214 atoms. This extended model contains second and third coordination sphere residues: Q406, H334, Y329, T580, G513, and additional H_2O that are located within hydrogen-bonding distance of R330, Y405, R330, H581 and H496, respectively (Fig. 2). The optimized structure reveals a five-coordinate square pyramidal geometry highly similar to **6** without significant distortion of any residues (Fig. 10a) relative to their crystallographic positions.

The Q406 residue that keeps the guanidium group of R330 in place was replaced by defining constraints for δN, εC, and χN atoms of R330. Importantly, the calculated electronic structure for the [Cu(II)–Y•] state is identical to model **6** as well (Fig. 10b). The antiferromagnetic coupling constant J was only slightly affected (approximately 608 cm^{-1}) with the same preference of the open-shell singlet ground state relative to the triplet state. Thus, model **6** provides a sufficient description of the structure of the [Cu(II)–Y•]. Model **7** may be needed in future investigation of the molecular mechanism of the GO chemical function.

Redox dependence of the active-site structure

While model **6** has been shown to be the converged computational model that reasonably reproduces key experimental data for the [Cu(II)–Y•] state of GO, the redox dependence of the active-site structure required a minor, yet important modification to this model. The change in the total charge of the computational model that accompanies reduction and the concomitant decrease of covalent interactions between the Cu center and its ligands required the introduction of additional geometric constraints to **6** that prevent unreasonable displacements of residues relative to their crystallographic positions. These additional constraints on β- and γ-C atoms of W290 affect the optimized model of the [Cu(II)–Y•] state negligibly. The introduction of these constraints is not arbitrary as they make up for the absence of hydrogen-bonding interactions from the Y329 residue that is present explicitly in model **7**.

The potential energy surface of the three redox states were rigorously probed as summarized in Fig. 3. In the following we will only discuss the lowest-energy, final stationary structures obtained for each redox step.

Semireduced state of GO [Cu(II)–Y]

Geometry optimizations from different starting points resulted in three different stationary points that are connected with a distortion coordinate along the Cu–O(Y495) and the Cu–OH_2 bonds. This distortion coordinate spans a tetragonal Cu(II) coordination geometry with short Cu–OH_2 (2.09 Å) and long Cu–O(Y495) (2.30 Å) distances, a structure with equivalent Cu–O distances (2.16 Å), and the lowest-energy semireduced [Cu(II)–Y] structure, which possesses a long Cu–OH_2 (2.24 Å) and a short Cu–O(Y495) (2.05 Å) bond (Table S1). The energy differences among these models are within the accuracy of the selected level of theory (approximately 4.4 kJ mol^{-1}). Going from the oxidized to the semireduced state, we find the most significant change is the shorter Cu–O(Y272) distance (2.02 vs. 1.92 Å) (Figs. 9a, 11). The calculated average Cu–O and Cu–N distances (1.99–2.05 and 2.00–2.05 Å)

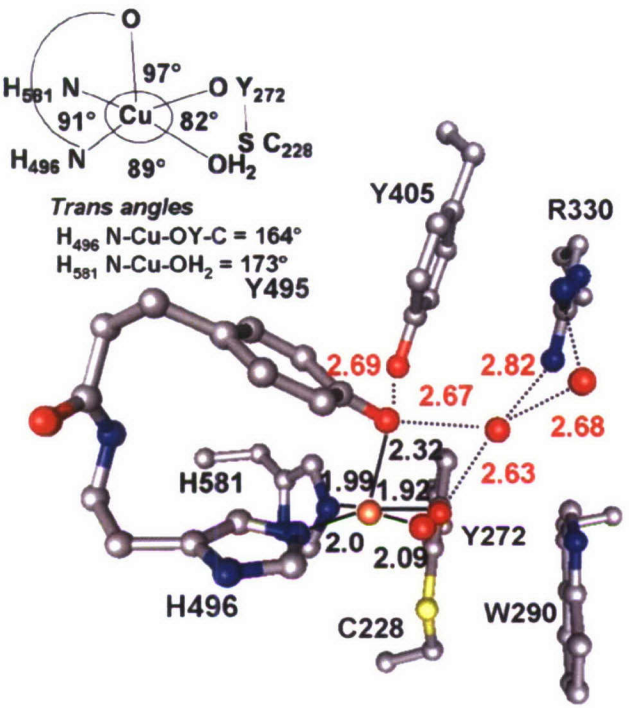


Fig. 11 Optimized structure of [Cu(II)–Y] from [Cu(II)–Y•] **6** with Cu–ligand (angstroms, *solid lines*) and hydrogen-bonding (angstroms, *dotted lines*) distances and ligand–Cu–ligand angles (degrees)

are slightly longer than the reported EXAFS [5] distances (1.97 and 1.95 Å). Independent of the position of a model structure along the distortion coordinate, the atomic spin densities remain similar: about 0.72–0.75 e^- on the Cu(II) ion, with the remaining 0.25 e^- distributed onto the equatorial ligands (Y272, 0.10–0.12 e^-; H496, H581, 0.04–0.07 e^-; H_2O, 0.02–0.04 e^-). TD-DFT analysis revealed that electronic transitions of these three structures are similar in energy to the experimental absorptions at 468 and 625 nm, which correspond to LMCT [Y272 (π) to σ^* of Cu(II) $3d_{x^2-y^2}$] and ligand-field transitions, respectively (Table S2).

The semireduced model (**7**) also parallels the result obtained from semireduced model **6** (Fig. S4, Table S3); however, the rotation of W290 is observed even in this extended model. Additional constraints (β-carbons of Y272 and W290) were applied to prevent this structural change, which had insignificant influence on the coordination geometry and the electron spin density distribution.

As a further validation of the semireduced computational model, we determined the one-electron redox potential for the [Cu(II)–Y•]/[Cu(II)–Y] couple. The ionization potential of [Cu(II)–Y•]/[Cu(II)–Y] is 7.88 V and by increasing the dielectric constant of the polarizable continuum surrounding model **6**, this decreases closer to the experimental value. The calculated redox potential is 691 mV with a dielectric constant of 10, which is a typical dielectric constant for a buried redox-active site (Table S4). However, the GO active site is partially solvent exposed and thus it is not surprising that a higher dielectric constant ($\varepsilon = 40$) gives a better agreement (calculated, 439 mV: experimental, 400 mV) [29]. The good agreement between the experimental and the calculated spectroscopic features, average Cu–O/Cu–N bond lengths, and redox potentials suggests that model **6** captures most of the important structural features and interactions in the protein environment for both the oxidized and the semireduced states of GO.

Reduced state of GO [Cu(I)–Y]

Starting from various initial structures for the reduced state (Fig. 3), we located stationary points along a similar distortion coordinate as described for the semireduced state. These range from a Cu(I) site with a trigonal pyramidal coordination geometry (higher energy) to the almost trigonal planar structure, where the Cu(I) is coordinated by two His N atoms and a Tyr O atom. The average Cu–(N/O) bond distance of 2.01 Å (Fig. 12a, Table S5) is close to the experimental EXAFS distance of 1.99 Å [6]. The π stacking of W290 and the Tyr–Cys cofactor is perturbed as expected from the increased electrostatic repulsion between the two-electron-reduced active site and the electron-rich W290. For model **7** only the trigonal planar structure (Table S6) was localized as a stationary point, with a W290 rotated away from the π-stacking plane owing to rotation along the βC–γC bond in the W290 residue.[1]

The redox potential calculated using a dielectric constant of 40 for the [Cu(II)–Y]/[Cu(I)–Y] couple is approximately 930 mV more negative than the experimental value of 160 mV [29]. This is a rather large difference and suggests that the reduced form may be protonated. Furthermore, Wright and Sykes [29] have described the pH dependence of the redox potential for the [Cu(II)–Y]/[Cu(I)–Y] couple. In order to investigate this, the reduced structure was protonated at O(Y495) and in a separate model at O(Y272). Both protonated states gave a three-coordinated Cu site (Fig. 12b, c), as also found by X-ray absorption spectroscopy [6], with Cu(I) ions coordinated by two N from His residues and O of tyrosinate, resulting in trigonal pyramidal structures. The redox

[1] It is important to note that the large number of constraints employed for the optimized structures for both model **6** and model **7** seem to have a negligible effect on the coordination geometry, yet rotation of W290 is observed for the reduced structure owing to the conformation flexibility of the βC–γC bond of W290. Since W290 is partially exposed to the solvent, it is quite likely that the distortion observed in the computational model can occur upon complete reduction of the GO active site in a protein environment.

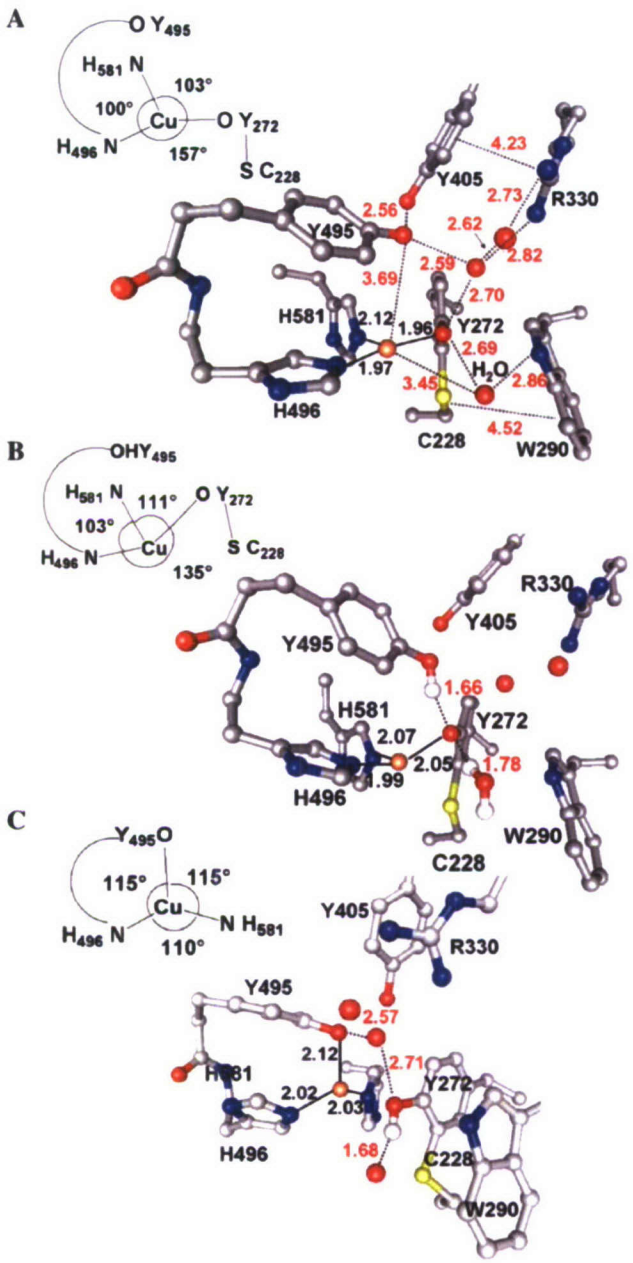


Fig. 12 Optimized structures of **a** [Cu(I)–Y], **b** Y495-protonated [Cu(I)–Y], and **c** Y272-protonated [Cu(I)–Y] with Cu–ligand (angstroms, *solid lines*) and hydrogen-bonding (angstroms, *dotted lines*) distances and ligand–Cu–ligand angles (degrees)

potentials for the semireduced/Y495-protonated and semireduced/Y272-protonated reduced couples in low dielectric medium ($\varepsilon = 10$) were calculated to be −163 and −40 mV, respectively, which are closer to the experimental value of 160 mV than was the redox potential of the nonprotonated, reduced form. A tetrasolvated hydronium (H_3O^+) ion was used as the source for the protons with an energetic cost of 11.9 eV.

Adjustment of the dielectric constant towards more polar environments ($\varepsilon = 40$) only slightly affected these redox potentials (−173 and −54 mV for Y495 and Y272 protonated states, respectively). While these potentials are somewhat less accurate than that of the oxidized/semireduced couple, the deviation of 300 mV between the calculated and the experimental potentials is already a rather reasonable agreement for a DFT computational model without the explicit consideration of point charges and dipoles from the rest of the protein environment.

Furthermore, the redox potentials calculated for a low dielectric medium ($\varepsilon = 10$) for the catalytically relevant two-electron redox couples of [Cu(II)–Y•]/[Cu(I)–HY495] and [Cu(II)–Y•]/[Cu(I)–HY272] are 264 and 325 mV, respectively. These values are in close vicinity of the experimental value (approximately 275 mV). In a more polar environment ($\varepsilon = 40$), these calculated redox potentials shift towards less positive values (133 and 193 mV) with a deviation of 80–140 mV from the experimental values.

Conclusion

It has already been well documented experimentally by site-directed mutations that residues of the second coordination shell influence substrate binding and catalysis in GO. Our current computational study parallels these site-directed studies in emphasizing the importance of the contributions from second coordination sphere residues. Furthermore, we would like to put this study forward as an example for evaluating contributions from the protein matrix to the active-site structure and catalytic activity. Inclusion of outer-sphere ligands in the *in silico* models was crucial to obtain the spectroscopically correct electronic structure for the active site of GO. The hydrogen-bonding interactions involving the O atom of Y495 and the π–π stacking of W290 and the Tyr•–Cys cofactor play a critical role in modulating the spin-density distribution around the Cu site and thus the molecular orbitals involved in catalysis. Since Y495 and W290 are both exposed to water solvent, inclusion of hydrogen-bonding interactions from explicit solvent water molecules with these residues is also essential. Our oxidized [Cu(II)-Y•] models (the smallest required model **6** and an optimal model **7**) with a tetragonal square-pyramidal active-site geometry and with a radical centered on the Tyr–Cys cofactor gives good agreement with the experimentally observed optical transitions and magnetic coupling constant. For an overview, the location of the protein radical, magnetic coupling and electronic transitions from all the computational models are provided as supplementary material (Table S7).

Using these models, we have developed the structure of the fully reduced [Cu(I)–Y] state, which is also implicated in the catalytic mechanism, as well as the resting

inactive, semireduced [Cu(II)–Y] state. The lowest-energy semireduced and fully reduced states were shown to have distorted tetragonal and trigonal planar Cu sites, respectively. We were able to estimate reasonably well the catalytically relevant redox potentials for the [Cu(II)–Y•]/[Cu(I)–HY] couple (133–193 mV; experimental, approximately 275 mV) in an intermediate dielectric environment with $\varepsilon = 40$. The protonation states of axial Y495 and the equatorial Y272 were found to be critical for obtaining experimentally sound redox potentials. Similarly to X-ray absorption spectroscopy studies on the substrate reduced form, both protonated states show a three-coordinate copper site. Overall model **6** is the smallest converged model for the GO active site which captures all the essential elements for the active site of GO and **7** is sufficiently extended to validate conclusions based on **6**. Our computational work sets the stage for a detailed investigation of substrate binding and the catalytic cycle of GO.

Acknowledgements We gratefully acknowledge the support of this research by NIH grant GM 27659 for D.M.D. and by the Montana State University Thermal Biology Institute (NASA NNG04GR46G) and the ONR Center for Bioinspired Nanomaterials (N00014-06-01-1016) for R.K.S.

References

1. Avigad G, Amaral D, Asensio C, Horecker BL (1962) J Biol Chem 237:2736–2743
2. Whittaker JW (2003) Chem Rev 103:2347–2363
3. Ito N, Phillips SEV, Stevens C, Ogel ZB, McPherson MJ, Keen JN, Yadav KDS, Knowles PF (1991) Nature 350:87–90
4. Whittaker MM, Whittaker JW (1988) J Biol Chem 263:6074–6080
5. Knowles PF, Brown RD, Koenig SH, Wang S, Scott RA, McGuirl MA, Brown DE, Dooley DM (1995) Inorg Chem 34:3895–3902
6. Clark K, Penner-Hahn JE, Whittaker M, Whittaker JW (1994) Biochemistry 33:12553–12557
7. Whittaker JW (2002) Adv Protein Chem 60:1–49
8. Whittaker MM, Whittaker JW (2000) Protein Expr Purif 20:105–111
9. Wachter RM, Branchaud BP (1998) Biochim Biophys Acta Protein Struct Mol Enzymol 1384:43–54
10. Rothlisberger U, Carloni P, Doclo K, Parrinello M (2000) J Biol Inorg Chem 5:236–250
11. Himo F, Eriksson LA, Maseras F, Siegbahn PEM (2000) J Am Chem Soc 122:8031–8036
12. Babcock GT, El-Deeb MK, Sandusky PO, Whittaker MM, Whittaker JW (1992) J Am Chem Soc 114:3727–3734
13. McGlashin ML, Eads DD, Spiro TG, Whittaker JW (1995) J Phys Chem 99:4918–4922
14. Gerfen GJ, Bellew BF, Griffin RG, Singel DJ, Ekberg CA, Whittaker JW (1996) J Phys Chem 100:16739–16748
15. Whittaker MM, Chuang YY, Whittaker JW (1993) J Am Chem Soc 115:10029–10035
16. Halfen JA, Young VG Jr, Tolman WB (1996) Angew Chem Int Ed Engl 35:1687–1690
17. Halfen JA, Jazdzewski BA, Mahapatra S, Berreau LM, Wilkinson EC, Que L Jr, Tolman WB (1997) J Am Chem Soc 119:8217–8227
18. Chaudhuri P, Hess M, Weyhermueller T, Wieghardt K (1999) Angew Chem Int Ed Engl 38:1095–1098
19. Chaudhuri P, Hess M, Mueller J, Hildenbrand K, Bill E, Weyhermueller T, Wieghardt K (1999) J Am Chem Soc 121:9599–9610
20. Chaudhuri P, Hess M, Flowers L, Wieghardt K (1998) Angew Chem Int Ed Engl 37:2217–2220
21. Kruse T, Weyhermuller T, Wieghardt K (2002) Inorg Chim Acta 331:81–89
22. Muller J, Weyhermuller T, Bill E, Hildebrandt P, Ould-Moussa L, Glaser T, Wieghardt K (1998) Angew Chem Int Ed Engl 37:616–619
23. Szilagyi RK, Solomon EI (2002) Curr Opin Chem Biol 6:250–258
24. Szilagyi RK, Metz M, Solomon EI (2002) J Phys Chem A 106:2994–3007
25. Szilagyi RK, Winslow M (2006) J Comput Chem 27:1384–1397
26. Ito N, Phillips SEV, Yadav KDS, Knowles PF (1994) J Mol Biol 238:794–814
27. Rogers MS, Knowles PF, Baron AJ, McPherson MJ, Dooley DM (1998) Inorg Chim Acta 275–276:175–181
28. Rokhsana D, Dooley DM, Szilagyi RK (2006) J Am Chem Soc 128:15550–15551
29. Wright C, Sykes AG (2001) J Inorg Biochem 85:237–243
30. Frisch MJ et al (2005) Gaussian 03, revision D 01. Gaussian, Pittsburgh
31. Schaefer A, Horn H, Alrichs R (1992) J Chem Phys 97:2571–2577
32. Rassolov VA, Pople JA, Ratner MA, Windus TL (1998) J Chem Phys 1223–1229
33. Francl MM, Hehre WJ, Binkley JS, Gordon MS, DeFrees DJ, Pople JA (1982) J Chem Phys 77:3654–3665
34. Hariharan PC, Pople JA (1973) Theor Chim Acta 28:213–222
35. Ryde U, Olsson MHM, Pierloot K (2001) Theor Comput Chem 9:1–55
36. Farkas O., Schlegel HB (1998) J Chem Phys 109:7100–7104
37. Solomon EI, Szilagyi RK, George SD, Basumallick L (2004) Chem Rev 104:419–458
38. Stratmann RE, Scuseria GE, Lovell T (1998) J Chem Phys 109:8218–8224
39. Davidson ER, Clark AE (2005) Int J Quantum Chem 103:1–9
40. Itoh S, Taki M, Kumei H, Takayama S, Nagatomo S, Kitagawa T, Sakurada N, Arakawa R, Fukuzumi S (2000) Inorg Chem 39:3708–3711
41. Rogers MS, Tyler EM, Akyumani N, Kurtis CR, Spooner RK, Deacon SE, Tamber S, Firbank SJ, Mahmoud K, Knowles PF, Phillips SE, McPherson MJ, Dooley DM (2007) Biochemistry 46:4606–4618
42. McGlashen ML, Eads DD, Spiro TG, Whittaker JW (1995) J Phys Chem 99:4918–4922
43. Whittaker MM, DeVito VL, Asher SA, Whittaker JW (1989) J Biol Chem 264:7104–7106

J|A|C|S
ARTICLES
Published on Web 03/07/2008

Dithiomethylether as a Ligand in the Hydrogenase H-Cluster

Arti S. Pandey, Travis V. Harris, Logan J. Giles, John W. Peters,* and Robert K. Szilagyi*

Department of Chemistry and Biochemistry and Astrobiology Biogeocatalysis Research Center, Montana State University, Bozeman, Montana 59717

Received December 17, 2007; E-mail: john.peters@chemistry.montana.edu; szilagyi@montana.edu

Abstract: An X-ray crystallographic refinement of the H-cluster of [FeFe]-hydrogenase from *Clostridium pasteurianum* has been carried out to close-to atomic resolution and is the highest resolution [FeFe]-hydrogenase presented to date. The 1.39 Å, anisotropically refined [FeFe]-hydrogenase structure provides a basis for examining the outstanding issue of the composition of the unique nonprotein dithiolate ligand of the H-cluster. In addition to influencing the electronic structure of the H-cluster, the composition of the ligand has mechanistic implications due to the potential of the bridge-head γ-group participating in proton transfer during catalysis. In this work, sequential density functional theory optimizations of the dithiolate ligand embedded in a 3.5−3.9 Å protein environment provide an unbiased approach to examining the most likely composition of the ligand. Structural, conformational, and energetic considerations indicate a preference for dithiomethylether as an H-cluster ligand and strongly disfavor the dithiomethylammonium as a catalytic base for hydrogen production.

Introduction

The [FeFe]- and [NiFe]-hydrogenases[1,2] catalyze the reversible hydrogen oxidation and proton reduction reactions. Crystal structures[3,4] of [FeFe]-hydrogenases reveal a biologically unprecedented active site (H-cluster) that exists as a [4Fe-4S] cubane linked to a 2Fe-subcluster via a cysteine thiolate. In addition, the 2Fe-subcluster is coordinated by terminal carbon monoxide ligands and a bridging carbon monoxide, terminal cyanide ligands, and a unique bridging dithiolate ligand. Although it has been generally accepted that the functional H-clusters from the structurally characterized [FeFe]-hydrogenases from *Clostridium pasteurianum*[3] and *Desulfovibrio desulfuricans*[4] are of the same composition, their structures differ by the coordination environment of the distal Fe atom of the 2Fe-subcluster by the presence of a water molecule. This difference has been attributed to the crystallization conditions and the likelihood that the enzymes in the two structures are poised in different oxidation states.[5]

An outstanding issue of considerable interest concerning the structure of the H-cluster is the chemical identity of the nonprotein dithiolate ligand. Because the H-cluster is a small component of the metalloprotein, attempts to unequivocally determine the composition of its dithiolate ligand have thus far been unsuccessful. However, this ligand has been of significant interest due to implications of its composition on the electronic structure of the H-cluster and the mechanism of dihydrogen uptake and evolution. In the first reported structure of an [FeFe]-hydrogenase isolated from *C. pasteurianum*,[3] the non-sulfur atoms of the dithiolate ligand were described as a covalent linkage of light atoms. In the characterization of the [FeFe]-hydrogenase from *D. desulfuricans*,[4] this ligand was originally assigned as propanedithiolate but later revised to dithiomethylamine.[6] Direct experimental evidence concerning the composition of the ligand has not yet been presented, and assignment of the secondary amine group in the ligand was based on its ability to serve as a proton donor or acceptor during catalysis. Regardless of the merits of the attributes of an amine as a potential catalytic base, it has been shown that the composition of the ligand has a large impact on the Fe−S bonding of the 2Fe-subcluster, which directly affects the redox active molecular orbitals of the H-cluster.[7] Thus, an imperative must be placed on resolving this interesting outstanding issue.

In this study, we undertook an unbiased investigation in which we considered the possibility of several dithiolate compositions with O, CH_2, NH, NH_2^+, and S groups in the γ-position (bridgehead) of the dithiolate ligand and evaluated their quantum chemically optimized structures in comparison to an experimental crystal structure of 1.39 Å resolution. We opted not to use an interfaced computational and a crystallographic code (COMQUM-X/Turbomole and CNS, respectively)[8] due to the current limitation of our generalized fragment method that allows us to obtain the correct electronic structure with spin

(1) Vignais, P. M.; Billoud, B.; Meyer, J. *FEMS Microbiol. Rev.* **2001**, *25* (4), 455−501.
(2) Adams, M. W. W. *Biochim. Biophys. Acta* **1990**, *1020* (2), 115−145.
(3) Peters, J. W.; Lanzilotta, W. N.; Lemon, B. J.; Seefeldt, L. C. *Science* **1998**, *282* (5395), 1853−1858.
(4) Nicolet, Y.; Piras, C.; Legrand, P.; Hatchikian, C. E.; Fontecilla-Camps, J. C. *Struct. Fold Des.* **1999**, *7* (1), 13−23.
(5) Nicolet, Y.; Lemon, B. J.; Fontecilla-Camps, J. C.; Peters, J. W. *Trends Biochem. Sci.* **2000**, *25* (3), 138−143.
(6) Nicolet, Y.; de Lacey, A. L.; Vernede, X.; Fernandez, V. M.; Hatchikian, E. C.; Fontecilla-Camps, J. C. *J. Am. Chem. Soc.* **2001**, *123* (8), 1596−1601.
(7) Schwab, D. E.; Tard, C.; Brecht, E.; Peters, J. W.; Pickett, C. J.; Szilagyi, R. K. *Chem. Commun.* **2006**, *35*, 3696−3698.
(8) Ryde, U. *Dalton Trans.* **2007**, *6*, 607−625.

10.1021/ja711187e CCC: $40.75 © 2008 American Chemical Society

Table 1. Data Collection Parameters

space group	$P4_22_12$
cell parameters	$a = b = 110.79$ Å $c = 103.57$ Å $\alpha = \beta = \gamma = 90.0°$
total observations	622 033
unique reflections	121 060

resolution shell (Å)	redundancy	completeness	avg $I/\sigma I$	R_{symm}[a]
46.9−4.8	7.8	98.2	31.5	0.06
4.80−3.40	8.7	99.9	31.1	0.07
3.40−2.77	7.9	100.0	24.7	0.09
2.77−2.40	6.6	99.9	20.2	0.09
2.40−2.15	5.7	99.7	17.2	0.10
2.15−1.96	5.1	99.5	14.4	0.12
1.96−1.81	4.6	98.9	11.5	0.14
1.81−1.70	4.2	97.4	8.6	0.17
1.70−1.60	3.9	93.5	6.2	0.22
1.60−1.52	3.9	89.0	4.5	0.29
1.52−1.45	4.1	84.3	3.1	0.41
1.45−1.39	4.4	81.7	2.5	0.57
Overall	5.1	93.5	12.6	0.09

[a] $R_{symm}(I) = \Sigma_{hkl} \Sigma_i | I_i - \langle I \rangle | / \Sigma_{hkl} \Sigma_i I_i$

polarized, open shell wave function. The latter is required for obtaining a proper magnetic coupling between the iron centers of the complete H-cluster. Instead of using structure factors from the crystallographic refinement, we present here a simple, yet insightful, new approach for evaluating various compositions for unknown atoms or groups. The density functional and basis sets chosen for the given study have already been thoroughly evaluated in a separate publication.[9] We present evidence for dithiomethylether as a bridging ligand of the 2Fe-subcluster. Furthermore, a conformation analysis of the orientation of the γ-group of the dithiolate indicates that dithiomethylamine would be unable to function as a catalytic base during catalysis. The presence of an ether functional group opens up new mechanistic possibilities to be explored by further synthetic and computational investigations.

Experimental Section

Crystallographic Refinement. The [FeFe]-hydrogenase in this study was purified from *C. pasteurianum* as described previously.[10] Crystals of the enzyme were obtained by the microcapillary batch diffusion method[11] in a precipitation solution of 25% polyethylene glycol 4000, 0.1 M sodium acetate (pH 4.6), and 0.1 M sodium sulfate as described previously.[3] All manipulations were conducted in an anaerobic chamber at room temperature. Crystals belonging to space group $P4_22_12$ were obtained in 7 days. Data was collected at SSRL on a Q4 detector at a wavelength of 0.95364 Å, processed with MOSFLM and scaled with SCALA.[12] The crystals were slightly sensitive to radiation damage during the relatively long exposure times (30 s to 1 min) that were necessary to collect the highest resolution data possible. The final data set consisted of a composite of data from four crystals and upon merging resulted in a data set with cumulative R_{symm} of 0.09 for data up to a resolution of 1.39 Å (Table 1).

Crystals were nearly isomorphous in comparison to the $P4_22_12$ crystals previously obtained with $a = b = 110.79$ Å, $c = 103.57$ Å, $\alpha = \beta = \gamma = 90°$.[13] For crystallographic refinement, a randomly selected R_{free} data set (3%) was assigned and the R_{free} was monitored through all stages of the refinement. Refinement was carried out with the program SHELX[14] using conjugate gradient least-squares[15] and block diagonal matrix least-squares methods. Protein coordinates and isotropic B factors were refined using conjugate least-squares method with protein bonds and angles restrained to target values from Engh and Huber.[16] Restraints for the iron-sulfur clusters were generated using SHELX-PRO.[14] Anisotropic B factors for all atoms were implemented for subsequent steps of refinement resulting in a drop of 3% in R_{free}. Coordinates were fitted to the *2Fo-Fc* and *Fo-Fc* maps in XTALVIEW[17] and alternate amino acid side chain conformations were added where applicable. Hydrogen atoms were added as fixed atoms. Another round of conjugated least-squares refinement was carried out with the bond distances for the H-cluster adjusted to those obtained from a DFT energy minimization that included the protein environment of the cluster within 3.5−3.9 Å. All reflections were included in the final step of the refinement resulting in an R_{cryst} of 13.7% for all reflections and 11.7% for reflections $>4\sigma$ (Table 3). All restraints were released for the calculation of standard uncertainties in bond lengths and angles of the clusters using blocked-matrix least-squares refinement. All main chain and side chain residues were included in the refinement with a limited number of residues modeled in multiple conformations and the final stereochemistry of the model was analyzed using PROCHECK.[18]

Computational Details. An approximately 200-atom virtual chemical model was constructed from the H-cluster and its 3.5−3.9 Å protein environment including the covalently bound full cysteine residues (italicized residues in Table 2). The dithiolate ligand is seated in a protein cavity that is lined with residues Cys299, Met497, Gly418, Phe417, Cys503 (Table 2) and capped with the 2Fe-subcluster. To avoid unreasonably large displacements relative to the close-to atomic-resolution crystal structure, the dithiolate ligand groups (bridge-head γ-group, β-methylene, μ-S) were allowed to relax only stepwise. All computations were carried out using Gaussian03 Rev. D.01.[19] We have selected the gradient-corrected Becke88[20] exchange and Perdew86[21] correlation functionals and the Stuttgart-Dresden effective core potentials and valence basis set for all atoms.[22] This level of theory has already been rigorously validated for iron-sulfur clusters in a previous work[9] and has been shown to reproduce iron-sulfur cluster geometry within 0.06 Å rms error. The [4Fe-4S] cubane of the H-cluster was treated as an open shell singlet structure in the +2 core charge state with two antiferromagnetically coupled rhombs of $M_s = \pm 9/2$. Using the 1.39 Å resolution structure of the H-cluster, 24 possible magnetic coupling schemes have been evaluated by our ionic fragment approach,[9] and only the lowest coupling was utilized in this study. The magnetic coupling schemes have also been evaluated with a more accurate functional and basis set (termed B5HFP86/BS5+ in ref 9) without observing any significant difference in the relative energies of magnetic states. The 2Fe-subcluster was described by a low spin $Fe^{(II)}/Fe^{(I)}$ pair in the $S = {}^1/_2$ state. We have considered all possible dithiolate compositions including the partially and fully protonated forms of the secondary amine group. For the sake of clarity of the plots and the discussion, the results obtained for the O, CH_2, and NH_2^- compositions are presented in the text and the results for the rest (S and NH in two different arrangements where the secondary amine group is H-bonding to the distal water and *vice versa*) are given as Supporting Information.

(9) Szilagyi, R. K.; Winslow, M. A. *J. Comput. Chem.* **2006**, *27* (12), 1385−1397.
(10) Chen, J. S.; Mortenson, L. E. *Biochim. Biophys. Acta* **1974**, *371* (2), 283−298.
(11) Georgiadis, M. M.; Komiya, H.; Chakrabarti, P.; Woo, D.; Kornuc, J. J.; Rees, D. C. *Science* **1992**, *257* (5077), 1653−1659.
(12) Collaborative Computational Project *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50* (Pt 5), 760−763.
(13) Peters, J. W.; Lanzilotta, W. N.; Lemon, B. J.; Seefeldt, L. C. *Science* **1998**, *282* (5395), 1853−1858.
(14) Sheldrick, G. M.; Schneider, T. R. *Methods Enzymol.* **1997**, *277*, 319−343.
(15) Konnert, J. H.; Hendrickson, W. A. *Acta Crystallogr.* **1980**, *A36* (May), 344−350.
(16) Engh, R. A.; Huber, R. *Acta Crystallogr.* **1991**, *A47* (4), 392−400.
(17) McRee, D. E. *J. Struct. Biol.* **1999**, *125* (2−3), 156−165.
(18) Laskowski, R. A.; McArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *265* (2), 283−291.
(19) Frisch, M. J.; et al. *Gaussian 03*, Rev. D01; Gaussian, Inc.: Wallingford, CT, 2006.
(20) Becke, A. D. *Phys. Rev. A: Gen. Phys.* **1988**, *38* (6), 3098−3100.
(21) Perdew, J. P. *Phys. Rev. B: Cond. Mater.* **1986**, *33* (12), 8822−8824.
(22) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86* (2), 866−872.

Table 2. Detailed List of Interactions between the Protein and the H-Cluster Considered for the Computational Model

residue[a]	molecular model	close contacts/interactions		
GLN195	H_2NCO	CO···S	3.33 Å	dipole interaction with S(Cys499)
HOH641	H_2O	O···S	3.00 Å	H-bonding to S(Cys355)
GLY302TRP303	H_2NCO	N···S	3.15 Å	H-bonding to S(Cys300)
CYS499	$^-SCH_2CH(CONH_2)NHCHO$	N···S	3.29 Å	H-bonding N(Ala498) to S([4Fe−4S])
		CO···S	4.04 Å	dipole interaction
THR356*CYS355*	$^-SCH_2CH(NHCHO)CONHCH_2CONH_2$	N···S	3.14 Å	H-bonding N(Ser357) to S(Cys355)
		CO···S	4.04 Å	dipole interaction
*CYS300*CYS299[b]	$^-SCH_2CH(CONH_2)NHCOCH_2CH_2SH$	CO···O	2.94 Å	H-bonding peptide to NH_3^+(Lys358)
				possible distal water-S(Cys299) interaction
CYS503[b]	$^-SCH_2CH(NHCHO)CONH_2$	bridging cysteine		
HOH17[b]	H_2O	O···N	3.11 Å	H-bonding to amide (Cys503)
MET497[b]	H_3CSCH_3	steric within 3.5 Å		bottom cap at the dithiolate ligand
GLY418[b]	H_2NCHO	CO···S	3.71 Å	dipole interaction with dithiolate
PHE417[b]	C_6H_6	C···O	3.21 Å	aromatic H-bonding to the distal water
LYS358	$H_3CNH_3^+$	N···N	2.74 Å	H-bonding to distal CN^-
MET353	H_3CSCH_3	S···O	3.15 Å	dipole with bridging CO
ALA320	H_3CCH_3	steric within 3.5 Å		top cap at bridging carbonyl
SER232	H_3COH	O···N	2.92 Å	H-bonding to proximal CN^-
PRO324	$[N(CHO)CH(CONH_2)C_3H_6]$	N···N	2.90 Å	H-bonding to distal CN^-
		steric within 3.5 Å		distal cap to the 2Fe-subcluster

[a] Ligands to H-cluster are italicized. [b] Residues around the dithiolate ligand.

Table 3. Progress of Refinement of the Previously Published Structure of CpI as Model against the 1.39 Å Data with SHELXL[a]

	resolution (Å)	*A*	*H*	*W*	Par	Obs	R_{cryst} % (all)	R_{free} % (all)	R_{cryst} % (>4σ)	R_{free} % (>4σ)
rigid body	2.5−46	4513	0	0	4522	22036	32.9	34.2	33.8	32.3
coordinates, isotropic B-factors	1.45−10	4844	0	331	19377	102188	21.9	25.4	20.0	23.5
addition of waters	1.45−10	4940	0	427	19761	102188	20.4	24.3	18.5	22.4
anisotropic B-values	1.45−10	4942	0	428	44477	102188	15.6	21.2	13.9	19.5
rebuilding + water + glycerol molecules	1.45−10	5208	0	632	45646	102188	13.9	19.9	12.3	18.2
include all data + anisotropic	1.39−10	5208	0	632	46871	113858	14.1	20.1	12.1	18.0
refine occupancies of split residues	1.39−10	5240	0	654	47045	113858	14.0	20.1	12.0	18.0
add Hs	1.39−10	5251	4300	667	47219	113858	13.2	19.4	11.8	17.8
final refinement	1.39−10	5251	4300	667	47219	119840	13.7	−	11.7	−

[a] All reflections were included during the final step of refinement. Columns include *A* - the number of non-hydrogen atoms, *H* - number of hydrogen atoms, *W* - number of water molecules, Par - number of parameters, and Obs - number of observations.

To evaluate the likelihood of the protonation, we carried out calculations for estimating the free energy of hydrolysis of the secondary amine bridge-head group of the dithiolate ligand. In addition to completing the first solvation shell of a solute (S) by hydrogen-bonded explicit water molecules, all solvation models were embedded in a low dielectric environment of $\epsilon = 10$ with solvent radius of 1.385 Å. Consequently, the structure of the S·n H_2O construct was fully optimized. Free energy corrections to the electronic energy were obtained from vibrational analysis and polarizable continuum calculations,[23] respectively. To validate this approximation, we also calculated the pK_a of a primary amine and methylthiolate at the same level of theory and employing the same solvation models that are in good agreement with common experimental values (10.7 and 7.5, respectively).

Although recently a more sophisticated method has been established for computational refinement[8] of protein crystal structures that takes into account the experimental structure factors we argue that gradually relaxed, partial geometry optimizations already have the potential to distinguish between various chemical compositions. Replacement of an ether group with a methylene or a secondary amine/ammonium group is expected to induce a considerable perturbation in the structure of the dithiolate ligand that can be quantified by comparing relative energies, initial gradients and forces, and the internal coordinates between the best crystallographic and the quantum chemistry optimized atomic positions. Naturally, a smaller energy change, smaller initial gradients and forces, as well as less deviation from the crystallographic positions would correspond to a more likely stoichiometry. We established a background for our approach by repeating the same set of optimization for two small molecule mimics $[Fe_2(CO)_6(pdt)]$[24] (**1**) and $[Fe_2(CO)_6(dtme)]$[25] (**2**), where pdt is propanadithiolate and dtme is dithiomethylether, as those used for the protein-embedded H-cluster in this paper. As a proof of concept, first both complexes **1** and **2** were fully optimized to obtain the most reasonable *in vacuo* structures (Figure S1A, top dotted lines and hollow circles for **1**, Supporting Information). Interchanging the bridge-head CH_2 and O groups in pdt and dtme containing complexes with O and CH_2, respectively, costs about 5 kJ/mol (stage A). Further optimization of the βCH_2 and the bridge-head γ-groups together (stage B) with the rest of the atomic positions kept fixed corresponds to about 31 kJ/mol further energy stabilization relative to the initial *in vacuo* optimized structures. Additional optimization of the bridging thiolate S atoms (stage C) contributes to the relaxation to a minor extent (2 kJ/mol) due to the geometric similarities of the $Fe_2(CO)_6S_2$ moieties (Figure S1B and C, Supporting Information) in complexes **1** and **2**. These clearly indicate that change in the chemical compositions of the bridge-head group should result in significant energy differences that are greater than the error limit of computations.

Interestingly, the same approach showed preference of the dithiomethylether composition for complexes **1** and **2** when the experimental

(23) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105* (8), 2999−3093.

(24) Lyon, E. J.; Georgakaki, I. P.; Reibenspies, J. H. Darensbourg, M. Y. *Angew. Chem., Int. Ed.* **1999**, *38* (21), 3178−3180.

(25) Song, L. C.; Yang, Z. Y.; Bian, H. Z.; Hu, Q. M. *Organometallics* **2004**, *23* (13), 3082−3084.

crystal structures were used as starting geometries (Figure S1, bottom dashed lines and half filled circles, Supporting Information). This seems to indicate that the dithiomethylether ligand has a less strained geometric structure to fit between the two iron ions in the $Fe_2(CO)_6$ fragment. However, this simplified evaluation needs to be taken with caution, because, as mentioned above for the *in vacuo* optimized structures, the dominant structural difference between the crystal structures of complexes **1** and **2** are not intrinsic to the small molecule complex (Figure S1B and C, Supporting Information), but they are due to the different supramolecular interactions[26] that give rise to crystal packing. To validate this, we carried out gradually relaxed, partial optimizations using a 225 atom containing fragment of the actual crystal structure of **1** (coordinates are given as Supporting Information). In this computational model, the bridge-head group (O or CH_2, stage A), then the bridge-head and βCH_2 groups (stage B), and last the full dithiolate ligand (stage C) of a central molecule were gradually relaxed (Figure S1A, middle solid lines and filled circles, Supporting Information). This central molecule was surrounded by eight identical molecules frozen in their crystal positions (Figure S1D, Supporting Information). This large model that closely resembles in size and complexity of the protein environment-embedded H-cluster now shows a preference of the CH_2 group over the O group in the pdt containing complex of **1** throughout the optimization. The smaller difference between the crystal packing embedded models relative to the *in vacuo* models is due to the weak intermolecular (dominantly dipole electrostatic) interactions between a central and its surrounding $Fe_2(CO)_6(SCH_2-CH_2/O-CH_2S)$ molecules.

It is also important to emphasize that in this study we have used a superior approach to the popular QM/MM approach by incorporating the complete 3.5−3.9 Å protein environment (see Table 2) into the computational model that is treated at quantum chemical level. We have already shown for a structurally less complex bioinorganic active site in galactose oxidase[27] than the H-cluster that this is crucial for reproducing experimental spectroscopic (UV−vis, magnetic coupling, ground state) and energetic (one and two electron redox potentials) parameters.

Results and Discussion

Crystallographic Refinement. In the current work, we have now refined the structure of the [FeFe]-hydrogenase from *C. pasteurianum* to 1.39 Å resolution. This improved resolution permits the inclusion of hydrogen atoms, the assignments of anisotropic *B*-factors, and calculation of individual atomic uncertainties in the range of those observed for small molecule mimics.[28−30] The final model exhibited comparable stereochemistry of the structure to that observed in the previous structure determined to 1.8 Å resolution with relative mean standard deviations from ideality in bond lengths and bond angles of 0.010 Å and 2.11°, respectively. A Ramachadran analysis of the structure indicated that all residues were found to be in either most favored (90%) or additionally allowed (10%) regions.[18] The overall *B*-factors for the protein were 17.1 $Å^3$, with main chain, side chains, iron sulfur clusters, and solvent atoms having *B*-factors of 14.0, 20.4, 9.5, and 33.7 $Å^3$, respectively. The current study is focused on the analysis of the H-cluster environment, and a more extensive analysis of high-resolution [FeFe]-hydrogenase structure is underway.

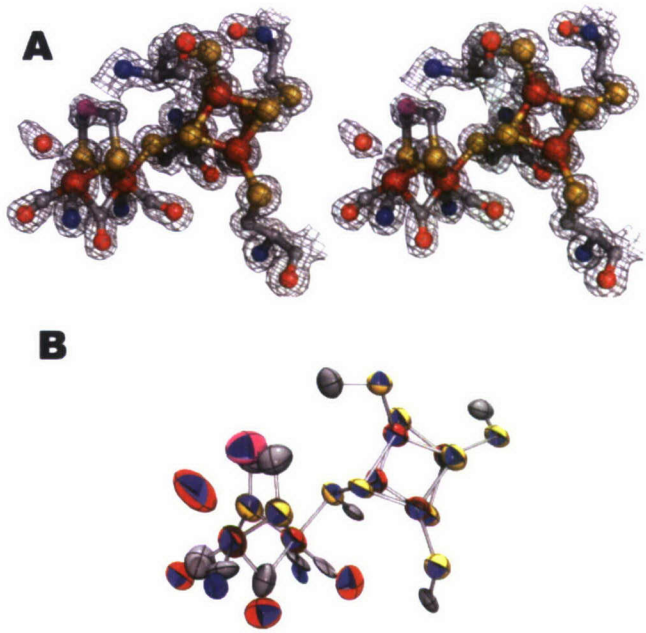
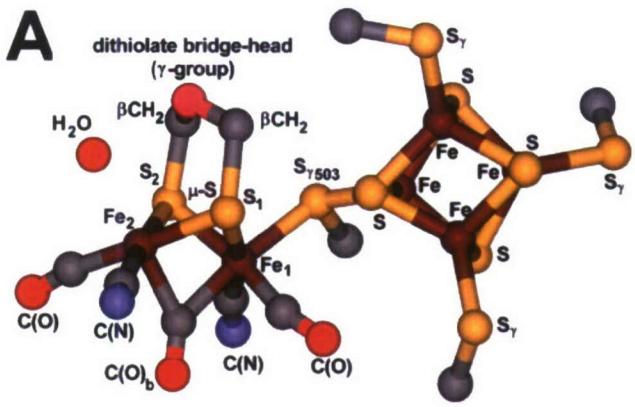


Figure 1. Stereoview of the H-cluster superimposed to 2*Fo*-*Fc* electron density map at contour level of 1.5σ (A); ORTEP presentation of the 2Fe cluster (B).

The electron density maps in this work are of excellent quality, and the majority of H-cluster atoms are resolved to atomic resolution (Figure 1A). The most characteristic bond lengths and intramolecular distances summarized in Figure 2 can be compared to small molecule mimics of the 2Fe-subcluster. The histograms of Fe···Fe (brown), Fe−S (orange), Fe−CO (red), and Fe−CN (blue) distances of selected 2Fe cluster mimics with both CO and CN ligands and a bridging thiolate were obtained from the Cambridge Crystallographic Database.[31] Although the average Fe···Fe distances in the small molecule mimics are highly similar (2.540 ± 0.030 Å) to those in the refined H-cluster (2.551 ± 0.007 Å, Figure 2). The different Fe oxidation states and the lack of the bridging CO ligand in the former are reflected in the slightly longer average Fe−S distances (2.288−2.320 ± 0.010 Å) of the H-cluster relative to the biomimetic compounds (2.278 ± 0.034 Å). However, the average Fe−CO and Fe−CN^- distances seem to agree rather well with those of the small 2Fe clusters (1.768 ± 0.024 and 1.925 ± 0.028 Å, respectively). The near-atomic resolution structure allows for the differentiation between the terminal CO and CN^- ligands, because the former are generally located closer to a metal center than the latter (see histograms in Figure 2). The refined internal coordinates highlighted in Figure 2 of the [4Fe-4S]-subcluster agree rather well with the analogous tetrathiolate coordinated [4Fe−4S] clusters (Fe···Fe 2.764 ± 0.100 Å, Fe−S(sulfide) 2.294 ± 0.041 Å, Fe−S(thiolate) 2.271 ± 0.047 Å, see distribution plots in Figure 2).

Despite the close-to atomic resolution of the current structure, the chemical composition of the dithiolate ligand of the H-cluster cannot be unambiguously assigned. Although the proximal (β-) groups of the dithiolate are rationally assumed to be methylene, there is still a debate on the composition of the central (γ-) dithiolate atom or group. Therefore, we carried out *in silico* model building and density functional theory calculations to

(26) Lehn, J. M. *Science* **1993**, *260* (5115), 1762−1763.
(27) Rokhsana, D.; Dooley, D. M.; Szilagyi, R. K. *J. Am. Chem. Soc.* **2006**, *128* (49), 15550−15551.
(28) Darensbourg, M. Y.; Lyon, E. J.; Zhao, X.; Georgakaki, I. P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (7), 3683−3688.
(29) Rauchfuss, T. B. Research on soluble Metal Sulfides: *Inorg. Chem.* **2004**, *43* (1), 14−26.
(30) Tard, C.; Liu, X. M.; Ibrahim, S. K.; Bruschi, M.; De Gioia, L.; Davies, S. C.; Yang, X.; Wang, L. S.; Sawers, G.; Pickett, C. J. *Nature* **2005**, *433* (7026), 610−613.
(31) Allen, F. H. *Acta Cryst.* **2002**, B58, (3, Sp. Issue 1), 380−388; (version 5.29; November 2007).



2Fe-subcluster	
$Fe_1 \ldots Fe_2$	2.551 ± 0.007
$Fe_2 \ldots OH_2$	2.379 ± 0.038
Fe_1-S_1	2.288 ± 0.010
Fe_1-S_2	2.300 ± 0.010
Fe_1-$S_{\gamma 503}$	2.391 ± 0.008
Fe_2-S_1	2.320 ± 0.010
Fe_2-S_2	2.300 ± 0.010
Fe_1-C(O)	1.740 ± 0.039
Fe_1-C(N)	1.825 ± 0.036
Fe_1-$C(O)_b$	1.906 ± 0.035
Fe_2-$C(O)_b$	1.991 ± 0.030
Fe_2-C(N)	1.929 ± 0.034
Fe_2-C(O)	1.765 ± 0.005
[4Fe-4S]-subcluster	
Fe-Fe	2.666-2.844 ± 0.006
Fe-S	2.231-2.387 ± 0.009
Fe-Sγ	2.273-2.303 ± 0.008

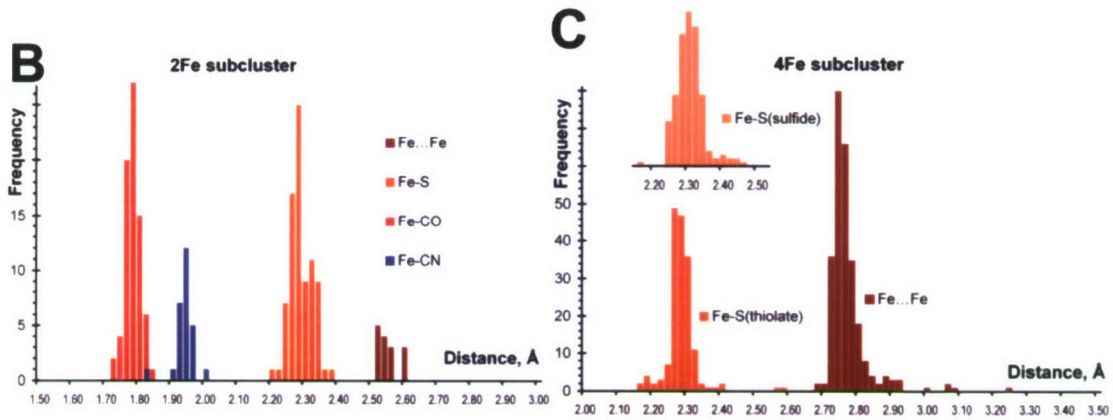


Figure 2. Selected interatomic distances within the H-cluster and atom labeling (A) adapted for the H-cluster. Histograms of Fe···Fe, Fe−S, Fe−CO, Fe−CN and Fe···Fe, Fe−S(sulfide), Fe−S(thiolate) distances within a selected set of 2Fe-subcluster mimics (B) and μ^3-tetrasulfido-tetrathiolato-tetrairon clusters (C).

evaluate the energetic and structural implications of various compositions (CH_2, NH, NH_2^+, O, and S) of the γ-group at the center of the dithiolate ligand.

As emphasized in the Computational Details, an accurate representation of the protein environment as a network of supramolecular interactions is essential for capturing the most important covalent, electrostatic, dipole, and hydrogen bonding interactions that all contribute to the structure and stability of the H-cluster. Table 2 summarizes the strongest intermolecular interactions identified in the 1.39 Å resolution structure between the H-cluster and its protein environment. As a result of this mapping, we found considerably more interactions for the 2Fe-subcluster than the 4Fe-subcluster of the H-cluster, which is expected due to the lack of a covalently bound residue to the 2Fe-subcluster. Notably, the greatest number of interactions (marked with superscript b in Table 2) involve the dithiolate ligand. It is important to mention that the dithiolate ligand in the protein bound H-cluster is involved in a more sophisticated and stronger network of intermolecular interactions than the small molecule mimics discussed in the Computational Details section (see above). The 4Fe-subcluster has only one H-bonding interaction to a sulfide and five dipole and/or H-bonding interactions to the thiolate groups of covalently bound cysteine residues. These interactions are expected to be influential in tuning the redox potentials of the 4Fe-sublcuster. About six important interactions surround the diatomic ligands in addition to residues Met353, Phe417, Ala320, and Pro324 lining the wall of the protein cavity around the 2Fe-subcluster. The position of the Cys299 residue is notable, because it is in H-bonding distance from the distal water and also connects the pool of water molecules adjacent to the distal Fe site (Fe_2) with the distal water.

Computational Evaluation of Various Dithiolate Compositions. Figure 3 presents the energy stabilization as first the γ-, next the β-, and last the μ-sulfur groups of the dithiolate ligand were gradually optimized while the positions of the rest of the atoms were kept fixed. In the first set of optimizations (section I), only the distal water was allowed to move. It was anticipated that the presence of the distal water could bias the optimizations due to hydrogen bonding interactions to the

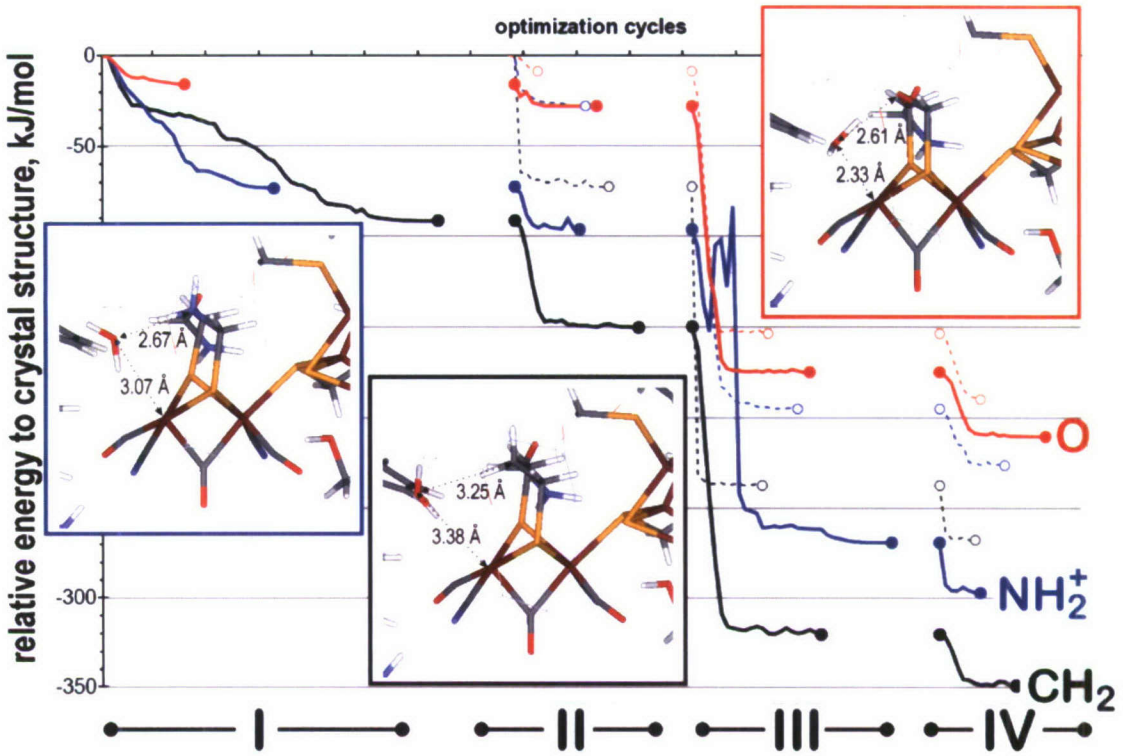


Figure 3. Energy stabilization upon sequential structural optimization of section I (distal water), II (water + γ-group), III (water + γ-group + β-CH_2), and IV (water + γ-group + β-CH_2 +μS) of the dithiolate ligand of the H-cluster embedded in a 3.5 Å protein environment. (Insets) Optimized cluster structures with selected intermolecular distances between the distal iron−distal water and bridge-head group−distal water.

γ-group. To test this, we also carried out optimizations in the absence of the distal water (dotted lines/hollow circles in Figure 3 starting in section II) in addition to optimizing the positions of the γ-group and the distal water together (not shown) without reaching a different conclusion regarding the preference for the γ-group composition. The changes in two characteristic distances are presented in Figure 4, and their optimized values are shown in the insets of Figure 3. The optimized coordinates of computational models are given as Supporting Information. For the dithiomethylammonium- and dithiomethylether-containing models, the distances of the distal water oxygen and the γ-group stay close to the experimental value of 2.60 Å. This distance is indicative of a strong hydrogen bonding interaction. From the structural insets in Figure 3, the hydrogen bonding interactions between the NH_2^+ or O groups and the distal water are actually different. In the former, the distal water is a hydrogen bond acceptor and donor in the latter. By comparing the water oxygen-distal iron distances, the dithiomethylether ligands remains close to the experimental value (2.33 Å vs 2.38 Å), whereas in the other cases, it moves away considerably from the distal Fe (CH_2: 3.38 Å and NH_2^+: 3.07 Å). Although the electron density feature and the thermal ellipsoid (Figure 1B) suggest dynamic flexibility and anisotropy for this water molecule, the crystallographic data would not support the assignment of the water molecule at the longer oxygen−iron distances observed for models with CH_2 and NH_2^+ groups. The energy change upon optimization is considerable less for the dithiomethylether ligand (O: −16 kJ/mol) than for the others (CH_2: −92 and NH_2^+: −73 kJ/mol) due to the good agreement between the optimized and experimental structures. Further optimization of the dithiolate γ- (section II, O: −28, CH_2: −150, NH_2^+: −96 kJ/mol), β-groups (section III, O: −175, CH_2: −321, NH_2^+: −269 kJ/mol), and last the thiolate sulfur atoms (section IV, O: −210, CH_2: −349, NH_2^+: −297 kJ/mol) clearly maintains the preference for an O linkage relative to CH_2 or NH_2^+ groups at the γ-position of the dithiolate ligand.

For the sake of clarity and simplicity, only the results obtained for the three dithiolate compositions with O, CH_2, and NH_2^+ γ-groups are presented in Figures 3 and 4. Figure S2 (Supporting Information) provides a more complete analysis of additional dithiolate γ-group composition including thioether (S) and nonprotonated amine (NH) as well as optimization results up to the entire 6Fe cluster embedded in the protein environment. The nonprotonated secondary amine group was examined in two different conformations in which the NH group acts as either a hydrogen bond donor (NH w/hydrogen bonding; cyan trace) or acceptor (NH w/o hydrogen bonding; teal trace). From the energetic and metric differences between the two protonated forms in Figure S2A−C (Supporting Information), the latter seems to be more favorable, yet it is less favorable than the dithiomethylether composition in all respects of the analysis. Comparing all the computational results, it is evident that the dithiomethylether composition corresponds to the least energy stabilization (Figure 2A) and the least deviations in internal coordinates (Figure S2B−C, Supporting Information) upon structural optimization relative to the 1.39 Å resolution structure of the H-cluster. It is also interesting to note that the distal water position is best reproduced if it is hydrogen bonding to the bridge-head γ-group. This is only possible if this group is either O or nonprotonated secondary amine.

It is expected that upon binding of a secondary amine substituted dithiolate to two positively charged transition metal ions the heat of hydrolysis of the amine group will decrease and thus the pK_a shift toward smaller values suggesting reduced

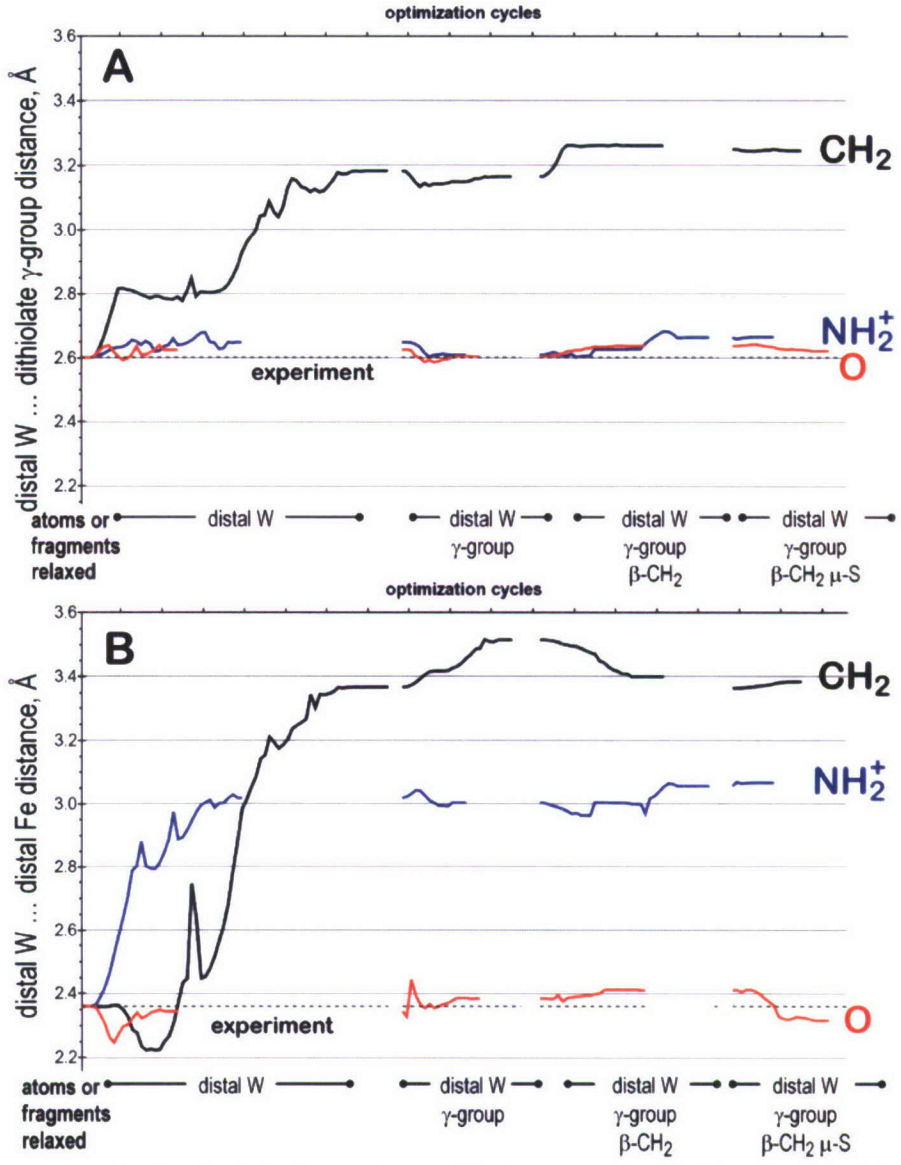


Figure 4. Interatomic distances as a function of optimization steps between dithiolate γ-group (A), distal iron-site (B), and the distal water (distal-W).

propensity for protonation. Using small molecule mimics of the 2Fe-subcluster ($Fe_2S_2(CO)_6$(dtma)), we estimated that the pK_a value of the dithiomethylamine in the most reduced Fe^IFe^I state to be about half (5.1) of a free secondary amine (Me_2NH, 10.1) in a simulated low dielectric environment ($\epsilon = 10$) with complete first aqueous solvation shell (see Computational Details). These calculations show that protonation of the secondary amine group in the most reduced Fe^IFe^I state is not likely.

Because a great deal of simplification had to be introduced in the above pK_a calculations, we further evaluated the conformational stability of the dithiolate ligand for a protonated versus nonprotonated ammonium group. Both [FeFe]-hydrogenase crystal structures show that the γ-group at the center of the dithiolate ligand is oriented toward the distal iron. The γ-group is not sterically constrained in the dimetalladithiabicyclohexane ring system; thus, it is fluxional under ambient conditions, as has been shown by detailed temperature-dependent NMR studies.[32] Partial structural optimizations of an alternate conformation for the γ-group in sections II, III, and IV of Figure 3 indicate that the conformation observed crystallographically is the most stable for all bridge-head group compositions except NH_2^+. In contrast, a 10−14 kJ/mol more-favored conformation of the dithiomethylammonium ligand exists with the secondary amine group pointing toward the thiolate S of Cys503 (Figure 5). The relative energies of these conformers remained the same when the model structures were fully optimized (Figure S2A − syn, dashed blue line; anti, solid blue line, Supporting Information). The energetically favored alternative conformation for the dithiomethylammonium ligand can be considered as a nonproductive catalytic base because it would shut down the proton shuttle to the distal iron site. Furthermore, it also competes with the proximal iron site of the 2Fe-sublcuters for the formally negatively charge bridging cysteine sulfur group that can compromise the structural integrity of the H-cluster framework.

Conclusions

The close-to atomic resolution crystal structure of the [FeFe]-hydrogenase from *C. pasteurianum* provided us a unique

(32) Lyon, E. J.; Georgakaki, I. P.; Reibenspies, J. H.; Darensburg, M. Y. *J. Am. Chem. Soc.* **2001**, *123* (14), 3268−3278.

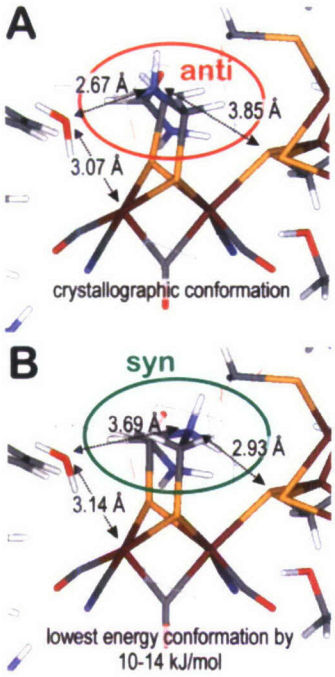


Figure 5. Anti (NH_2^- points away from the 4Fe-subcluster) and syn (NH_2^- points toward the 4Fe-subcluster) conformation of the dithiomethylammonium ligand.

opportunity for critically evaluating the composition of the dithiolate ligand by an unbiased systematic *in silico* analysis. Computations carried out for a more than 200 atom model of the H-cluster and its immediate 3.5−3.9 Å protein environment favor the dithiomethylether composition. Our estimated pK_a value of the $Fe^{I}Fe^{I}$-bound dithiomethylamine suggests that amine group is not likely protonated under physiologically relevant conditions. In addition, conformational analysis of the dithiolate ligand disfavors the dithiomethylammonium as the catalytic base for hydrogen production due to a preferred alternate conformation for the secondary ammonium group. This conformation has not been observed even in the lower resolution crystal structures. In this alternate conformation, the positively charged ammonium group hydrogen bonds to the sulfur atom of the bridging cysteine and thus effectively shuts down the catalytically important electron-delocalization between the 4Fe- and the 2Fe-subclusters of the H-cluster. Oxygen as the central atom has not yet been critically exploited by synthetic and computational biomimetic chemistry; however, our results provide a support for this as the next generation of targets of H-cluster models.

Acknowledgment. Portions of this research were carried out at the SSRL operated by Stanford University on behalf of the DOE BES. The SSRL Structural Molecular Biology Program is supported by the DOE, OBER, and by the NIH, NCRR, Biomedical Technology Program, and the NIGMS. This research was supported by the NASA Astrobiology Biogeocatalysis Center at Montana State University funded by NASA (NNA08CN85A). T.V.H. and L.J.G. acknowledge the undergraduate research support from MSU Center for Bio-Inspired Nanomaterials (ONR N00014-06-01-1016).

Supporting Information Available: Full optimizations profiles for all dithiolate compositions, validation of the computational analysis, coordinates of computational models, and complete ref 19. This material is available free of charge via the Internet at http://pubs.acs.org. The FeFe-hydrogenase coordinates have been deposited to the Protein Databank (PDB ID 3C8Y).

JA711187E