LA-UR-99- 2839

96612

Title: **Landscapes and Dynamics of Proteins**

Author(s):	Niels Anderson	University of Washington
	Joel Berendzen	LANL, P-21
	Stefan Boettcher	Emory University
	Mark Buchanan	Nature Magazine
	Kelvin Chu	University of Vermont
	Chuali Chang	University of California, Irvine
	Richard Dyer	LANL, CST-4
	Doyne J. Farmer	The Prediction Company
	Hans Frauenfelder	LANL, CNLS, P.I.*
	Klaus-Dieter Fritsch	University of Bayreuth
	Angel Garcia	LANL, T-10
	Robert Gennis	Univ Ill at Urbana- Champaign
	Klaus Gerwert	Ruhl Univ, Bochum, Germany
	Maximo Granados	Univ Autonoma de Zacatecas
	Lydia Gregoret	Univ of California, Santa Cruz
	Peter Hanggi	University of Augsburg
	Bai-lin Hao	Chinese Academy of Sciences
	Gerhard Hummer	LANL, T-10
	Eric Jeckelmann	University of California, Irvine
	Ioannis Kevrekidis	Princeton University
	Janos Ladik	Univ of Erlangen-Nuremberg
	Gustavo Lopez	Univ of Guadalajara, Mexico
	Ben McMahon	LANL, CNLS
	M. Muthukumar	University of Massachusetts
	Ulrich G. Nienhaus	University of Ulm, Germany
	Robert Nigmatulin	Rensselaer Polytechnic Institute
	Eric Olson	Indiana University
	Pal Ormos	Biophys Inst, Szeged, Hungary
	Graham Palmer	Rice University

Fritz Parak	Tech Univ, Munich, Germany
Matteo Pellegrini	Univ of Calif, Los Angeles
Alan Perelson	LANL, T-10
Michel Peyrard	Ecole Normale Superi de Lyon
Cong-Kha Pham	Tokyo University
George Pimbley	Self-employed
David Pines	Univ Ill at Urbana-Champaign
Anders Sandvik	University of Illinois
Rudolf Rigler	Karolinska, Institute
Michael Schick	University of Washington
Joerg Schmalian	University of Illinois
Gunther Schoeck	University of Vienna
Alfons Schulte	University of Central Florida
Solomon Stavrov	Tel Aviv University
Daan Thorn-Leeson	LANL, CNLS
Zoltan Toroczka	VA Polytech Inst./State Univ
Peter Wolynes	Univ Ill at Urbana-Champaign
William Woodruff	LANL, CST-4

Submitted to: DOE Office of Scientific and Technical Information (OSTI)

Los Alamos
NATIONAL LABORATORY

Form 836 (10/96)
ST 2629

Landscapes and Dynamics of Proteins

Authors

Niels Anderson	University of Washington
Joel Berendzen	LANL, P-21
Stefan Boettcher	Emory University
Mark Buchanan	Nature Magazine
Kelvin Chu	University of Vermont
Chuali Chang	University of California, Irvine
Richard Dyer	LANL, CST-4
Doyne J. Farmer	The Prediction Company
Hans Frauenfelder	LANL, CNLS,*(e-mail: frauenfelder@lanl.gov)
Klaus-Dieter Fritsch	University of Bayreuth
Angel Garcia	LANL, T-10
Robert Gennis	University of Illinois at Urbana-Champaign
Klaus Gerwert	Ruhl University, Bochum, Germany
Maximo Granados	University Autonoma de Zacatecas
Lydia Gregoret	University of California, Santa Cruz
Peter Hanggi	University of Augsburg
Bai-lin Hao	Chinese Academy of Sciences
Gerhard Hummer	LANL, T-10
Eric Jeckelmann	University of California, Irvine
Ioannis Kevrekidis	Princeton University
Janos Ladik	University of Erlangen-Nuremberg
Gustavo Lopez	University of Guadalajara, Mexico
Ben McMahon	LANL, CNLS
M. Muthukumar	University of Massachusetts
Ulrich G. Nienhaus	University of Ulm, Germany
Robert Nigmatulin	Rensselaer Polytechnic Institute
Eric Olson	Indiana University
Pal Ormos	Bisophysics Institute, Szeged, Hungary
Graham Palmer	Rice University
Fritz Parak	Technical University, Munich, Germany
Matteo Pellegrini	University of California, Los Angeles
Alan Perelson	LANL, T-10
Michel Peyrard	Ecole Normale Superieure de Lyon
Cong-Kha Pham	Tokyo University
George Pimbley	Self-employed
David Pines	University of Illinois at Urbana-Champaign
Anders Sandvik	University of Illinois
Rudolf Rigler	Karolinska, Institute
Michael Schick	University of Washington
Joerg Schmalian	University of Illinois
Gunther Schoeck	University of Vienna
Alfons Schulte	University of Central Florida
Solomon Stavrov	Tel Aviv University
Daan Thorn-Leeson	LANL, CNLS
Zoltan Toroczka	Virginia Polytechnic Inst. & state University
Peter Wolynes	University of Illinois at Urbana-Champaign
William Woodruff	LANL, CST-4

Abstract

This is the final report of a three-year, Laboratory Directed Research and Development (LDRD) project at Los Alamos National Laboratory (LANL). Complex systems have become a major focus for research in physics, chemistry, and materials science. Four aspects are essential for understanding complex systems: structure, energy landscape, dynamics, and function. Proteins, the building blocks of living systems, are ideal for studying the properties and laws of complex systems. We have been exploring heme proteins, in particular myoglobin, experimentally, theoretically, and computationally. We have arrived at a model that combines structure, energy landscape, dynamics, and function for an apparently simple biological process: the binding of a small molecule to a protein. The model can most likely be generalized to many other proteins in general and to other complex systems. The understanding may lead to novel proteins, novel materials, and possibly also to new insight into problems such as biological threats. The enormous complexity of biological systems means that the extension to other systems is not a short-term endeavor, but must be continued for a considerable time to reap the full benefit of such studies.

Background and Research Objectives

Complex systems, from glasses to biomolecules and the immune system, all appear to share one major property: Their ground state is highly degenerate and must be described by an energy landscape (EL). The more complex the system, the more involved is the EL. In "simple" systems, from atoms to nuclei, and crystalline solids, studies of the energy levels have been crucial for a full understanding of the physics. We expect that similar progress will be made on complex systems by studying the EL. At present, however, even in the least complex systems such as glasses and spin glasses, the EL is not fully known. Even less is known about the laws that govern the dynamics in complex systems, in other words, the motions in the hyperspace of the EL. Some general approximate laws have been found, but a deep understanding is still lacking. The goal of this work is to investigate, in close contact with experiment, the EL and the relevant reaction both theoretically and computationally.

In addition to energy landscapes we also propose to study various types of fitness landscapes. One type of fitness landscape is the affinity landscape where to each antibody

---

*Principal Investigator, e-mail: frauenfelder@lanl.gov

gene sequence an affinity for a particular ligand is assigned. Equivalently, the antibody can be fixed and a set of ligands with differing sequences examined. In either case, depending upon the rule of assigning affinity to sequence, different landscapes result. We will examine the properties of different landscapes such as the number of local optima and the statistical properties of move-rules on a landscape. For instance, if only uphill moves are allowed, what is the probability distribution for the number of mutations until one gets trapped at a local optimum? How does the dynamics change if the landscape has large flat regions that arise due to the many-to-one nature of the mapping between sequence and structure? We shall also examine the effects of correlation among the affinities of different sequences on the ability to search sequence space for high-affinity molecules. Such searches can be conducted in the laboratory by a biotechnology company or in vivo by the immune system during the course of affinity maturation. We believe that elucidating the search strategies employed by the immune system may give insights into strategies that biotechnology companies can exploit.

The structure of a molecule influences its fitness. An interesting example of this effect is seen in RNA viruses. If we view the rate of replication of the virus as its fitness, then one can show that the fitness depends on the structure of the RNA molecule. We can begin to model this effect by using RNA secondary-structure-prediction algorithms to assign a secondary structure to each sequence, and then assign a fitness based on the degree of base pairing in the structure. We can then ask questions about the evolution of the viral sequence based on mutation rate and rules for fitness based-selection.

Applied molecular evolution, or molecular diversity, is a general new approach to the discovery of useful molecules and to the analysis of structure-function relationships in such molecules. Applications range from discovery of new pharmaceuticals, to biosensors, to the generation of materials with novel properties. The central idea is to generate vast libraries for potentially useful molecules. Thereafter, successive rounds of mutation, recombination, and selection are employed to yield increasingly improved variants in a manner somewhat akin to natural evolutionary processes. The "landscape" of the property of interest, e.g. binding affinity, is therefore sampled at a variety of points in a structured way.

The research objectives follow logically from the problems just sketched. In a few selected systems, structure, energy landscape, dynamics, and function must be studied in detail and the connections must be established. At first, this path appears to be straightforward. The work performed under this LDRD project and research performed during the past 25 years proves, however, that even apparently simple biological processes are sophisticated and beautifully executed.

Importance to LANL's Science and Technology Base and National R&D Needs

Our work contributes directly to two DOE Competencies:

Advanced Computing, Modeling, and Simulation

Bioscience and Biotechnology.

More specifically, the work performed during the past three years and the work that we hope to perform in the future can have the following impact:

1. It will help establish a science base for work on biological threats.
2. At present, most of the work in the biosciences and biotechnology is done with limited input from the physical and computational sciences. Our research constructs a rigorous base for the design and function of biopolymers.
3. Young scientists working on the interface between physics, chemistry, computing, and biology, as in our research, will form the cadre for the future DOE efforts in bioscience and biotechnology.
4. The DOE has started a new initiative on "Complex and Collective Phenomena". Our work, and related efforts in other groups at LANL, will give us a head start in contributing to this initiative.
5. The biosciences and biotechnology will be an increasingly important part of the work at LANL, from efforts on bio-threats, to structural and functional genomics. At present, much of this work at LANL and in other laboratories is done without strong interdisciplinary connections linking biology, chemistry, computing, and physics. Our work is fully integrated and can be a model for future major initiatives.

Scientific Approaches and Accomplishments

We describe here the major accomplishments. Where the results have been published, we will be brief. We provide details only on work that has not yet been published. The accomplishments fall into three categories, research, workshops and conferences, and lectures.

Research-The Approach. (For details see Frauenfelder, LA-UR 95-4357. The Complexity of Proteins.) Myoglobin (Mb), our "hydrogen atom of biology", is a moderately sized protein that contains a heme group with a central iron atom. This central iron atom binds dioxygen and carbon monoxide. In the simplest version, the reaction with CO can be written as

$$Mb + CO \rightarrow MbCO.$$

Initially, this reaction was considered to be a one-step process. Our work, starting before 1975, showed that the process is far more complicated.[1] By performing experiments over a wide range of time and temperature (from about 10 to 300K) we found that the reaction occurs in a series of steps. Despite hundreds of experimental, computational, and theoretical publications, the process has not been fully understood until now.

The approach to study the Mb reaction is in principle very simple. MbCO is placed into a cryostat. A short laser pulse breaks the Mb-CO bond and CO moves away from the iron. The subsequent rebinding, Mb+CO $\rightarrow$ MbCO is observed optically. The reaction pathway is deduced from the time and temperature dependence of this process.

Why is this reaction important? How is the problem related to the energy landscape and dynamics? What can we learn that is relevant for other proteins and other complex systems? What could the impact of a complete solution be for bioscience and biotechnology? The answers to these questions are not fully known, but some aspects are already clear.

Research – Importance. At present, the function of not even one protein is understood quantitatively. Understanding one protein fully would help understand others and possibly lead to the construction of mutated proteins with desired properties, a goal important for biomaterials, medicine, and pharmacology. If myoglobin were a member of an unimportant

class, a complete understanding would not be too exciting. Mb, however, is a member of the hemoglobin family.[2] Figure 1, copied from Hardison[2], shows that the hemoglobins appear extremely widely and can have very different functions. Understanding Mb quantitatively can be the key to many secrets.

Research – Structure, Energy Landscape, Dynamics, and Function. For many years, proteins were considered to be nearly static structures. While a few papers pointed to the importance of motions, the beautiful structures obtained by x-ray diffraction gave the impression of well-ordered systems, with a unique structure and a well-defined energy. The work in Ref. 1 proved that this impression was wrong; a given protein can assume a very large number of somewhat different conformations, called conformational substates (CS).[3] The energy landscape describes the connection between protein structure and energy. Protein motions (protein dynamics) are described as jumps between various CS. The goal of our work is to connect structure, energy landscape, dynamics, and function for the case of the binding of CO to Mb. The problems to be solved are the following. After photodissociation, the CO leaves the Mb with a high probability. When CO tries to enter again, it has to compete with the water molecules that surround the protein. How can it compete effectively? What are the protein motions that are involved? A complete answer to these questions will be given in a paper in preparation, but some of the essential features are sketched here.

A major step towards the goal of connecting structure and function was made by the work of Joel Berendzen and his collaborators.[4] They measured the structure of the Mb-CO complex at low temperature before and after photolysis Figure 2 shows the region around the heme in Mb. In Figure 2a, the CO is bound to the heme iron; in Figure 2b, the CO sits in the heme pocket parallel to the heme. Figure 2 also shows the existence of two pockets, space not occupied by atoms of the protein. The so-called heme pocket is on the side where the CO is bound (distalside), the so-called xenon pocket is on the opposite (proximal) side.

Both pockets are formed mainly by residues that have been conserved in the course of evolution.

The exit and entry of CO can now be described as follows. Immediately after photodissociation, the heme iron moves partially out of the heme plane, making rebinding of the CO difficult. For a few femtoseconds, the CO ricochets around the heme pocket and then docks at the site shown in Figure 2b. After about 100 nanseconds, it moves out into the solvent, either directly or through the xenon pocket. When there is no CO in the heme pocket, a water molecule can move in. Because there are many more water than CO molecules around the protein, the entrance of water is much more likely and water thus can block the direct reentrance of the CO into the heme pocket. The CO therefore moves to the xenon pocket and waits there until the water moves out of the heme pocket. It then transits to the distal side and binds. This scenario connects structure, energy landscape, dynamics, and function. The quantitative description in terms of these features is, of course, more intricate and we are in the course of writing a complete version. It is intriguing to note that xenon pockets have been found in many other proteins. Do these have an important, as yet unrecognized, function?

Workshops and Conferences. The Center for Nonlinear Studies (CNLS) organizes a large number of workshops and conferences each year. Some of the ones that were held during the past three years were directly connected with the subject of this LDRD project. These conferences establish contacts between LANL researchers and an outstanding group of U. S. and foreign scientists. At the same time, they provide visibility to the LANL research. We list here the two most important conferences.

1. Landscape Paradigms in Physics and Biology: Concepts, Structures, and Dynamics, 16th Annual International Conference of the Center for Nonlinear Studies, Los Alamos, NM, May 13-17, 1996. This conference attracted an outstanding group of researchers from a wide range of fields. The papers from the conference appeared in Physica D 107, 117-435 (1997) and are collected in a volume edited by H. Frauenfelder, A. R. Bishop, A. Garcia, A. Perelson, P. Schuster, D. Sherrington, and P. J. Swart and published by Elsevier in 1997.

2. Third International Symposium on Biological Physics, Santa Fe, NM, September 20-24, 1998. This symposium, organized mainly by CNLS, was very well attended. A volume is in preparation.

Lecture Series and Lectures. A lecture series on "Biology for Physicists" exposed many of the LANL physicists to the basic biological concepts. In addition, Hans Frauenfelder presented a large number of lectures, describing the CNLS work, both at national and international conferences.

Publications

1. Frauenfelder, H. "Proteins and the Physics of Complexity," in Nonlinear Excitation in Biomolecules, Pp 177-189, M. Peyrard, Ed., Springer-Verlag, Berlin (1995).

2. Frauenfelder, H. "Komplexitat in der belebten Materie" Schriftenrcihe Ernst-Abbe-Kolloquium Jena. 9, 1-26 (1995).

3. Frauenfelder, H. "Proteins: A Challenging Many Body Problem," Phil. Mag. B74, 579-585 (1996).

4. Johnson, J. B., Lamb, D. C., Frauenfelder, H., Muller, J. D., McMahon, Ben, Nienhaus, G. U. and Young, R. D. "Ligand Binding to Heme Proteins VI: Interconversion of Taxonomic Substates in Carbomonoxymyoglobin," Biophys. J. **71**, 1563-1573.

5. Frauenfelder, H. "The Complexity of Proteins, In Physics of Biological Systems-From Molecules to Species," Eds. H. Flyvbjerg, J. Hertz, M. H. Jensen, O. G. Mouritse, and K. Sneppen, Springer-Verlag, Verlin, Pp. 29-60 (1997).

6. Frauenfelder, H. "Proteins as Paradigms of Complex Systems," in Rontgen Centennial, Eds. A. Haasc, G. Landwehr, and F. Umbach, World Scientific, Singapore, Pg. 227-241 (1997).

7. Frauenfelder, H. and Wolynes, P. G., "Picturing the Working Protein," in Theoretical and Computational Genome Research, S. Suhai, Ed., Plenum Press, New York, Pp 231-236 (1997).

8. Frauenfelder, H. "Proteins and Glasses," in Structure and Dynamics of Glasses and Glass Formers, C. A. Angell, K. L. Ngai, K. Kieffer, T. Egami, and G. U. Nienhaus, Eds., Materials Research Society, Pittsburgh, Pp. 343-347, (1997).

9. Garcia, A. F., Krumhansl, J. A. and Frauenfelder, H., "Variations on a Theme by Debye and Waller: From Simple Crystals to Proteins," Proteins, 29, Pp.153-160 (1997).

10. Nienhaus, G. U., Muller, J. D. McMahon, B. H. and Frauenfelder, H. "Exploring the Conformational Energy Landscape of Proteins," Physica D 107, 297-331 (1997).

11. Frauenfelder, H. McMahon, B. Stojkovic, B. P. and Chu, K., "Is Myoglobin like a Swiss Watch?," in Nonlinear Cooperative Phenomena in Biological Systems, L. Matson, Ed., World Scientific, Singapore, pp.145-164 (1998).

12. Frauenfelder, H, "Protein Dynamics and Function," in Protein Dynamics, Function, and Design, O. Jardetzky and J. F. Lefevre, Eds., Plenum Press, New York, pp-95-102 (1998).

13. Frauenfelder, H. and McMahon, B., "Dynamics and Function of Proteins," Proc. Natl, Acad. Sci. USA, 95, 4795-4797 (1998).

14. Frauenfelder, H and Thorn Leeson, D., "The Energy Landscape in Non-biological and Biological Molecules," Nature Structural Biology, 5, 757-759 (1998).

15. Deissler, Robert J. and Brand, Helmut R., "The Effect of Nonlinear Gradient Terms on Breathing Localized Solutions in the Quintic Complex Ginzburg-Landau Equation," Physical Review Letters, **81** pp 3856-3859 (1998).

16. Frauenfelder, H., Wolynes, P. G., and Austin, R. H., "Biological Physics," Rev. Mod. Phys., Centennial Issue 1999 (in press).

References

[1] Austin, R.H., Beeson, K., Eisensein, L., Frauenfelder, H., and Gunsalus, I.C., "Dynamics of Ligand Binding to Myoglobin," Biochemistry **14**, 5355-5373 (1975).

[2] Hardison, R. "The Evolution of Hemoglobin," American Scientist **87**, 126-137 (1999).

[3] Frauenfelder, H., Petsko, G.A., and Tsernoglou, D., "Temperature-Dependent X-Ray Diffraction as a Probe of Protein Structural Dynamics," Nature **280**, 558-563 (1979).

[4] Schlichting, I., Berendzen, J., Phillips Jr., G.N., and Sweet, R.M., "Crystal Structure of Photolysed Carbonmonoxy-Myoglobin," Nature **371**, 808-812 (1994).

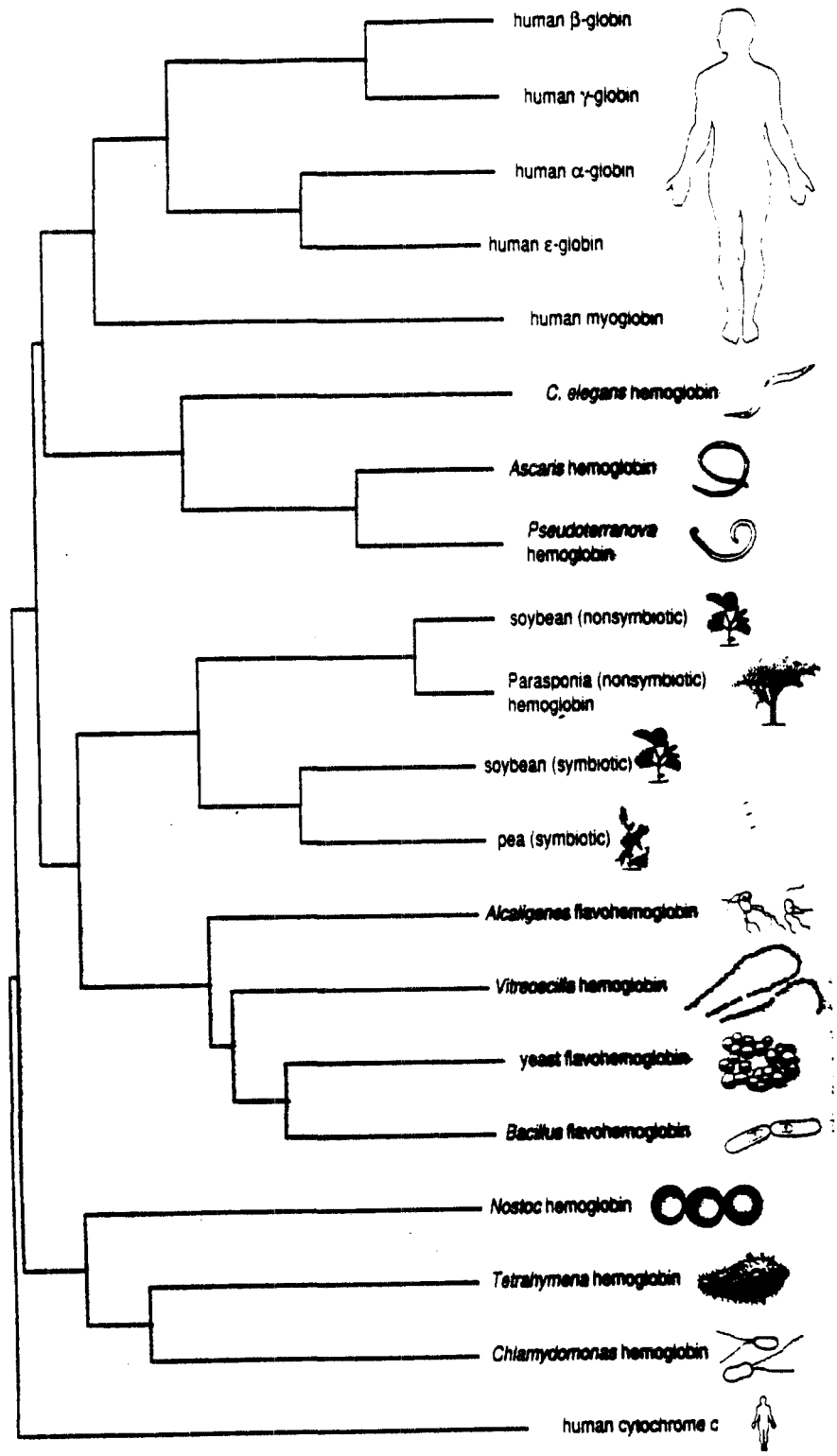

Fig. 1 The family of hemeglobins. (From Hardison, ref.2.)

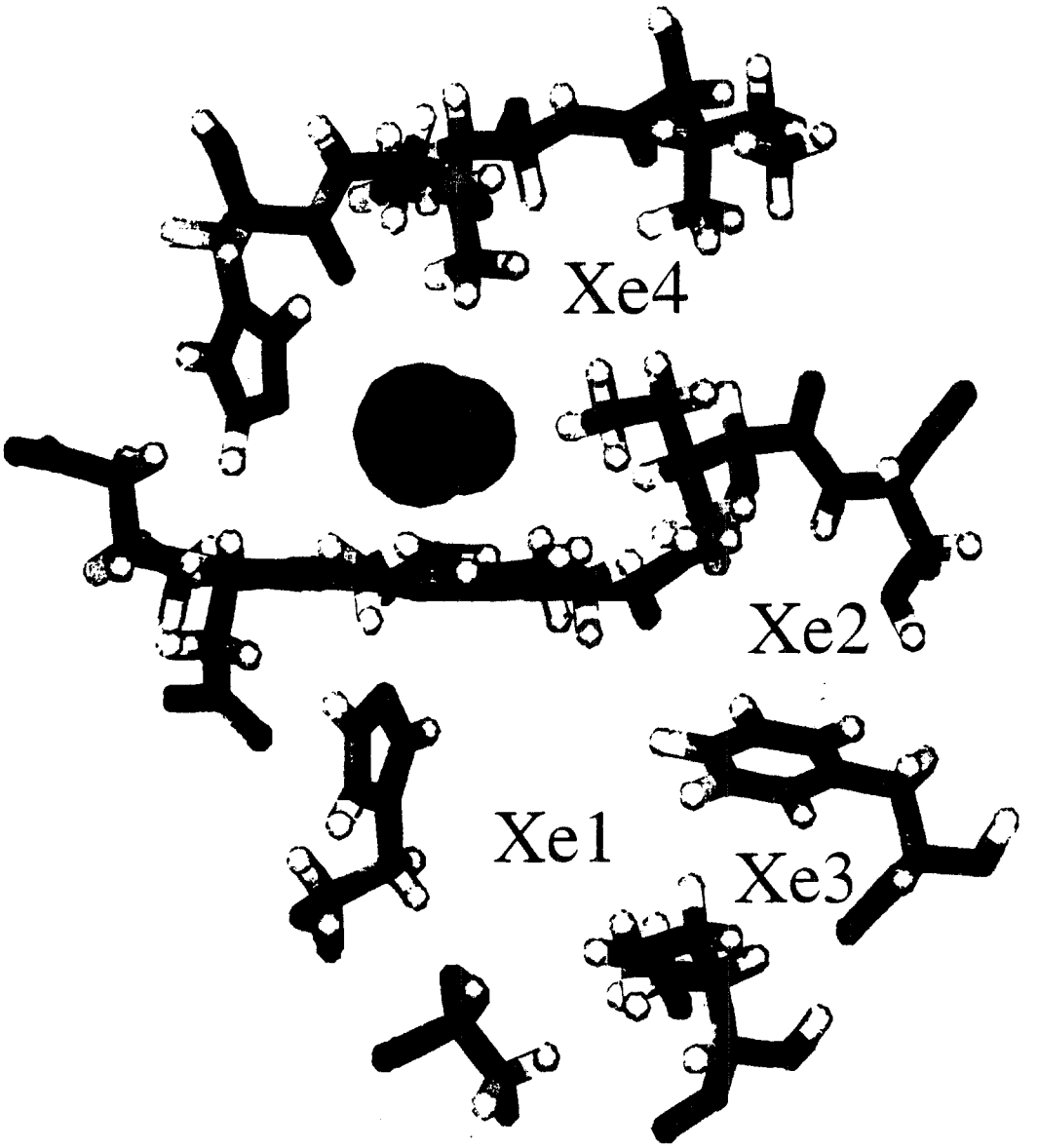

Fig. 2 A cross section through the interior myoglobin. a. The bound state, MbCO.b. The photodissociated state. (Courtesy J. Berendzen, LANL P 21)