Title: | **Biological Neutron Scattering: Now and the Future**

Author(s): | **Jill Trewhella, CST-4**

Submitted to: | **International Workshop on Science in Neutron-Arena of JHP
KEK, Tsukuba, Japan
March 26-27, 1996**

# Los Alamos
## NATIONAL LABORATORY

**MASTER**

Form No. 836 R5
ST 2629 10/91

## BIOLOGICAL NEUTRON SCATTERING:  NOW AND THE FUTURE

Jill Trewhella

Chemical Science and Technology Division, Los Alamos National Laboratory,
Mail Stop G758, Los Alamos NM 87545

ABSTRACT

Neutrons have an important role to play in structural biology.  Neutron crystallography, small-angle neutron scattering and inelastic neutron scattering techniques can all contribute unique information on biomolecular structures.  In particular, solution scattering techniques can give critical information on the conformations and dispositions of the components of complex assemblies under a wide variety of relevant conditions.  The power of these methods are demonstrated by examples protein/DNA complexes, and $Ca^{2+}$-binding proteins complexed with their regulatory targets.  In addition, we demonstrate the utility of a new structural approach using neutron resonance scattering.  The impact of biological neutron scattering to date has been constrained principally by the available fluxes at neutron sources and the true potential of these approaches will only be realized with the development of new more powerful neutron sources.

## 1. Introduction

The 20th century has seen great advances in medicine and biotechnology that have been founded on detailed knowledge of the structures and interactions of biological polymers:  the polynucleotides, DNA and RNA, which contain coded information that determines all the molecular structures needed to assemble a living system; and the polypeptides, or proteins, that carry out almost all the "work" of living systems and are the ultimate expression of the coded information carried in the DNA.  It has been the miracle of nature to design these polymers using unique linear sequences of a relatively small number of chemical subunits (4 nucleotide acids for DNA, 20 amino acids for proteins) such that they form highly specialized three-dimensional structures that fold reliably into their functional forms. Understanding biological function has become a study of these molecular forms, in sufficient detail to be able to describe and predict the physics and chemistry they perform in the microenvironments they create by bringing together specific chemical subunits.  The latter half of the 20th century has seen rapid advances in technologies that have led to an impressive structural data base on biological polymers.  Importantly, molecular biology and cloning techniques has enabled structural biologists to routinely obtain large amounts of pure samples for physical studies.  X-ray crystallography has provided a wealth of high resolution structural information on proteins and polynucleotides that can be coaxed to form crystals.  NMR spectroscopy with isotope labeling has opened the window to high resolution structure determination of small proteins, and even small polynucleotides in solution.  Optical and laser spectroscopy techniques are used to probe the active sites of proteins at high resolution, and provide critical information on protein structural dynamics over the many orders of magnitude relevant to biological function (femtoseconds to hours).  Electron microscopy and a number of novel new microscopy techniques are giving clearer images of molecular assemblies.

In this world of high technology structural tools, neutron techniques have a special place.  As non-ionizing radiation, neutrons are benign probes of biomolecular structures, that can locate

hydrogen atoms the very high precision in crystals, give information on the shapes and dispositions of the components of biomolecular complexes in solution, as well as probe the molecular dynamics of biological molecules over very broad range of timescales. My own work using neutrons has focused on studies of biomolecular complexes in solution using small-angle scattering.

## 2. Small-Angle Neutron Scattering

Small-angle scattering from monodisperse particles in solution gives information on the particle shape. With neutrons, one can capitalize on the large difference in neutron scattering amplitude between the isotopes of hydrogen ($^1H$ and $^2H$). Selective deuteration of one component of a complex of proteins in solution provides a way of altering the mean neutron scattering density of that component. Further, by changing the deuterium level in the solvent, the neutron scattering contrast of each component is varied. If internal scattering density fluctuations are negligible, then "solvent matching" can be achieved by adjusting the deuteration level in the solvent so that the mean solvent and particle (or component) scattering densities are the same, i.e. there is zero contrast and hence no small-angle scattering from the particle rendering it "invisible" in the neutron experiment. "Contrast variation" techniques provide methods for extracting structural information on the individual components of a complex and their relative dispositions [1,2]. In the case of DNA/protein complexes one can take advantage of the natural difference in neutron scattering density between DNA and protein, thus making contrast variation experiments feasible without labeling either component. This ability to study interacting components in a complex or assembly is extremely important if you wish to understand biomolecular function, since healthy function depends upon highly coordinated interactions between biomolecules that allow for signals to be transmitted in order to obtain the appropriate molecular responses to physiological stimuli. Highlighted here are a few examples from my own work that demonstrate the power of neutron small-angle scattering from biomolecules in solution.

## 3. Studies of Gene 5 protein/DNA Complexes

Essential proteins that bind nonspecifically, cooperatively, and preferentially to single stranded DNA (ssDNA) are required in the life cycles of viruses, prokaryots, and eukaryots. The ssDNA binding protein (the gene 5 protein, or g5p) of filamentous bacterial viruses is among the smallest of such proteins, having a molecular weight just under 10,000 daltons, and there is a wealth of detailed information about the physical properties of g5p which is widely regarded as a "model" ssDNA binding protein whose further study will elucidate fundamental mechanisms governing this class of protein/DNA interactions. The g5p forms a dimer (Fig. 1), and electron microscopy of fd g5p complexed with fd-ssDNA shows left-handed nucleoprotein superhelices with variable pitch.

We undertook small-angle scattering experiments in combination with electron microscopy to provide definitive experimental data to test models for g5p/fd-ssDNA superhelices, as well as to evaluate how unique the models might be [3]. In addition, we obtained new information on the dependence of nucleotide type and protein/nucleotide (P/N) ratio on the structure of the complexes. Reconstituted complexes were made with single-stranded fd viral DNA (fd ssDNA), poly[d(A)] and poly[r(A)]. All complexes formed similar left-handed, flexible superhelices (Fig. 2) having approximately the same diameter, but the pitch differs among these complexes. The g5p dimers associate to form a superhelical framework to which the polynucleotide is attached. A combination of X-ray and neutron small-angle scattering data was needed to prove the nucleic acid is inside the protein superhelix. The neutron and X-ray experiments provided two sets of scattering data in which the relative contributions of the protein and DNA to the scattering were significantly different ("contrast variation"). These data,

along with known structural parameters, provided sufficient constraints to define the relative locations of the DNA and protein. A Monte Carlo integration modeling procedure applied to the scattering data was used to systematically test large numbers of possible superhelical models for each complex, and previously proposed models based on parameters obtained from electron microscopy were found to be essentially correct and unique. The data on the complexes with different P/N ratios showed that mass per unit length values decreased while the rise per dimer and pitch of the superhelix increased for g5p/fd-ssDNA complexes with decreasing P/N ratios.
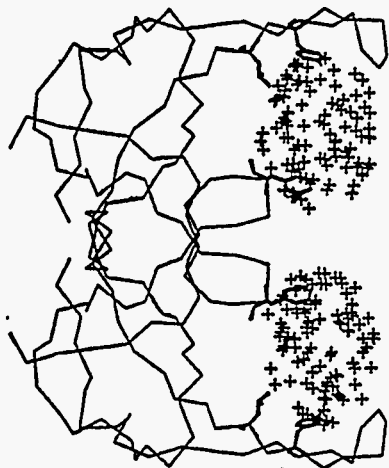


*Fig. 1:  g5p dimer with the its two ssDNA-binding sites depicted by crosses. The protein structure is depicted as a trace of its α-carbon backbone.*
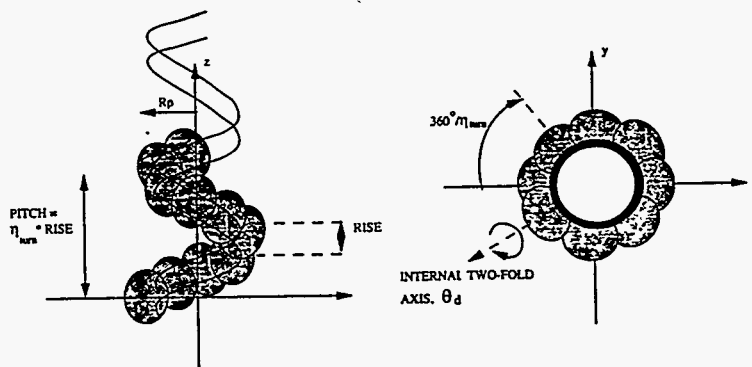
*Fig. 2:  Schematic drawing of the superhelix construction for the g5p/DNA complex. Ellipses represent the g5p dimers, the 2 ssDNA's interior to the protein envelop are indicated as solid lines.*

## 4.   Protein Complexes in Biochemical Regulation

The divalent calcium ion is the simplest of natures "messengers" used to regulate a broad range of cellular processes. $Ca^{2+}$ messages are generally "delivered" when $Ca^{2+}$ binds to a protein and triggers a cascade of molecular interactions leading to a physiological response. Small-angle neutron scattering has proven extremely useful for studying the evolutionarily related, structurally similar dumbbell-shaped $Ca^{2+}$-binding proteins calmodulin and troponin C and their interactions with the target proteins whose activity they regulate.

*4. 1  Calmodulin Target Enzyme Interactions*
Calmodulin is a multifunctional protein that acts as the major intra-cellular receptor for $Ca^{2+}$, and regulates the activity of a diverse array of target enzymes including a large number of kinases. The crystal structure of calmodulin shows it has two globular domains connected by an extended, solvent exposed α-helix of 7-8 turns [4] (Fig. 3). Each globular domain is made up of two pairs of helices that form a "cup-shape", with 2 $Ca^{2+}$-binding sites at the base of the cup (Fig. 5). $Ca^{2+}$ binding to calmodulin causes the pairs of helices to move away from each other exposing hydrophobic residues lining the inner surface of the cup. Small-angle neutron (and X-ray) scattering experiments showed calmodulin contracts about target enzyme binding domains with the common characteristic of having a high propensity for forming a basic, amphipathic α-helix (for review see [5]). Fig. 3 shows the structure of calmodulin complexed with its helical target binding domain from myosin light chain kinase, determined at high resolution by NMR [6] and crystallography [7] subsequent to our neutron scattering experiments [8]. The contraction of calmodulin is achieved via flexibility in the interconnecting helix region of the

binding to different arrangements of hydrophobic and charged residues important in forming these complexes.
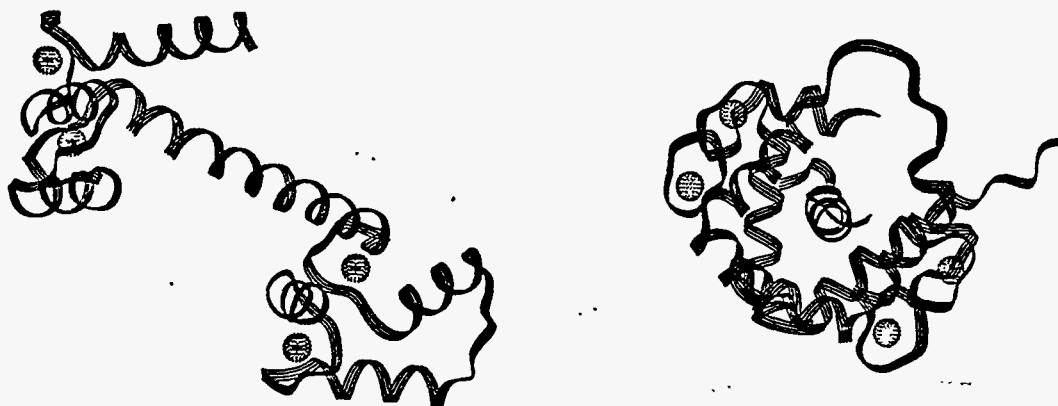


*Fig. 3: α-carbon backbone trace showing the structures of 4Ca²⁺•calmodulin (left) and 4Ca²⁺•calmodulin complexed with its binding domain in myosin light chain kinase (right).*

A second type of calmodulin binding domain has been identified in the catalytic subunit of phosphorylase kinase. This system is distinctive because calmodulin is part of a multisubunit enzyme and its binding domain encompasses approximately 70 residues (compared with the more usual 17-25 residues). Using deuterium labeling and neutron scattering we showed calmodulin remains extended in its interaction with the catalytic subunit of phosphorylase kinase and the nature of the interaction is fundamentally different to the previously characterized contracted structures [9]. The ability of calmodulin to modulate its conformation via flexibility in its interconnecting helix region in order to accommodate different target binding domains is a remarkable example nature building functional diversity as well as specificity into a compact, unusual shape.

### 4.2 Troponin C and Troponin I

Like calmodulin in phosphorylase kinase, troponin C (TnC) is an integral component of a multisubunit complex: troponin. Muscle contraction/relaxation is achieved when thick and thin filaments made up of complex assemblies of proteins slide past each other. The sliding mechanism is regulated by $Ca^{2+}$ binding to TnC, resulting in a conformational change that transmits a signal to troponin I (TnI), which in turn releases its inhibition of the interaction between the thick and thin filaments which drives the sliding action. Our neutron scattering experiments show that, similar to calmodulin in phosphorylase kinase, TnC remains extended in its complex with its regulatory target TnI. The neutron data further show that TnI is even more extended than TnC and the centers-of-mass of the two components are approximately coincident [10]. We have modeled the complex [11] with TnC having a structure very similar to its crystal structure [12,13] and TnI forming a spiral structure that encompasses the TnC (Fig. 4). The TnI spiral has the approximate dimensions of an α-helix and its central portion winds through the two hydrophobic cup-shaped regions in each of TnC's two globular domains.

This solution structure of 4Ca²⁺•TnC•TnI gives important insights into the molecular basis for the $Ca^{2+}$-dependent regulation of muscle contraction. In the $Ca^{2+}$-bound complex, both ends of the TnI central spiral region are anchored by interactions with the hydrophobic cup regions of TnC, and the TnI inhibitory sequence is constrained to be associated with TnC and the inhibitory function is switched off. The loss of $Ca^{2+}$ from the N-terminal low-affinity $Ca^{2+}$-specific binding sites results in a closing of that domain, lowering its affinity for TnI at the N-

terminal end, and allowing the TnI inhibitory sequence (96-115) the flexibility to shift from its binding site on TnC to that on actin, switching the inhibition on. The regulatory signal is thus transmitted via the central spiral region of TnI, which contains the inhibitory sequence.
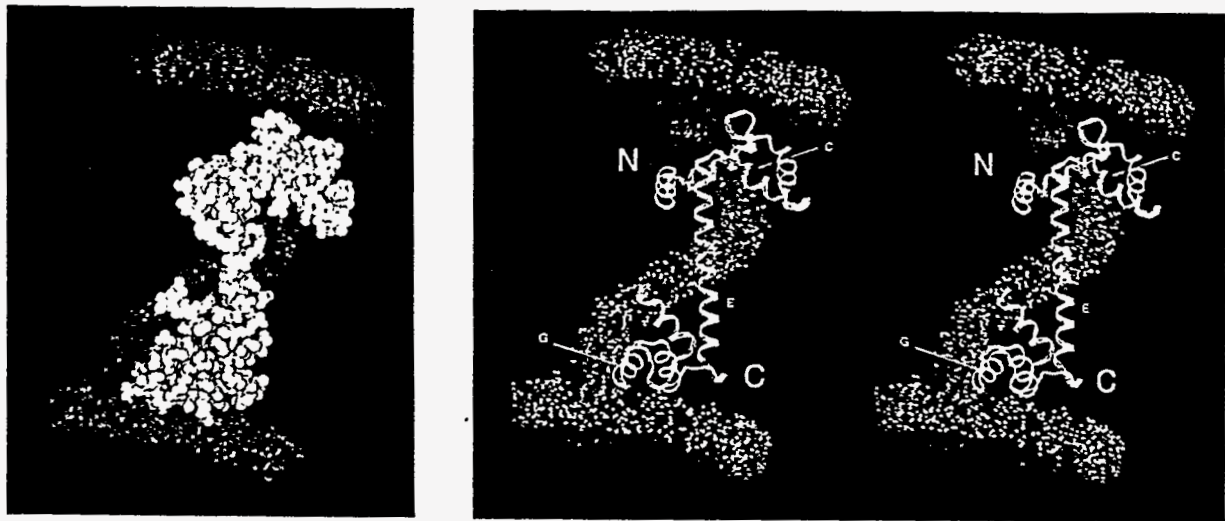


*Fig. 4: Right Stereo drawing of the model of 4Ca2+•TnC•TnI. The white crosses depict TnI, which follows a spiral path winding around the TnC represented as an α-carbon backbone ribbon drawing. The N- and C-terminal domains of TnC are labeled, as are helices C, E, and G which were known to interact with TnI. Left The same model with TnC represented as a CPK model.*

## 5. Neutron Resonance Scattering from $^{240}$Pu-labeled Calmodulin

Nuclei with strong resonances in their coherent neutron scattering amplitudes include the soft metals $^{113}$Cd, $^{115}$In, rare earths $^{149}$Sm, $^{151}$Eu, $^{157}$Gd, $^{163}$Dy, $^{168}$Yb, $^{177}$Hf, the noble metal $^{193}$Ir, and actinides $^{230}$Th, $^{240}$Pu [14,15]. Of these nuclei, $^{240}$Pu has the strongest and sharpest resonance (at 0.28 Å, Fig. 6) that lies within the wavelength regime accessible using the Low-Q
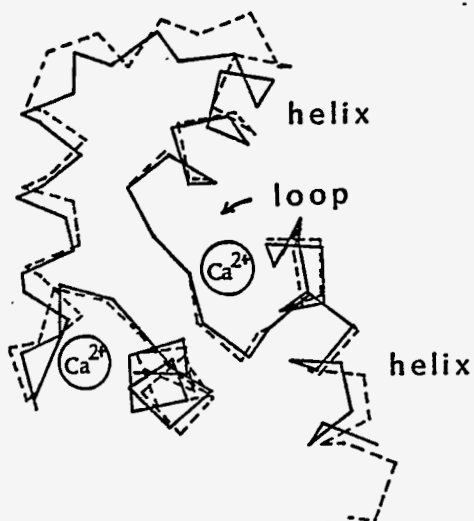


*Fig. 5: Backbone structures of the structurally homologous N- and C-terminal domains of calmodulin showing the $Ca^{2+}$-binding sites that were labeled with $^{240}$Pu.*
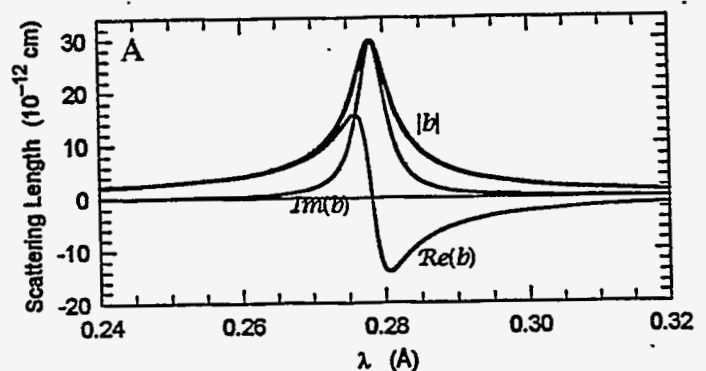
*Fig 6: Coherent scattering length for $^{240}$Pu in the vicinity of the resonance.*

diffractometer (LQD) at the Manuel Lujan Jr Neutron Scattering Center (MLNSC), a pulsed neutron source at Los Alamos. A number of the resonant nuclei have properties that are similar to biologically relevant metals. For example, $^{240}Pu^{3+}$ is a positively charged ion with the same ionic radius as $Ca^{2+}$. We therefore undertook experiments in which we sought to substitute $^{240}Pu$ for $Ca^{2+}$ in calmodulin (Fig. 5) with the intent of measuring the resonance scattering form the $^{240}Pu$ ions to determine distance information. The experiment was initially aimed at demonstrating the utility of this technique as a structural tool, and in the process we also demonstrated that $^{240}Pu$ binds specifically, with high affinity to the $Ca^{2+}$ sites in the protein [16]. We were successful in extracting the resonance scattering data, and were able to determine the average distance between $Ca^{2+}$-binding sites in each globular domain of calmodulin as 11.83 $\pm$ 0.42 Å. This was in excellent agreement with the values determined from the crystal structure coordinates.

## 6. Conclusions

With the powerful capabilities neutrons have for probing structure and dynamics in biological molecules, one might ask why are neutron scattering techniques not more widely used? And why are there not more examples of high profile results from neutron scattering studies of biological systems? The answer is simple. There are a very limited number of neutron sources with the needed instrumentation and user support, and the neutron fluxes available are at best marginal for biological studies. Neutrons interact weakly with biological systems, the samples are generally difficult to prepare in large quantities, particularly when isotopic labels are called for. Spallation neutron sources are a number of orders of magnitude from their potential intensities and may be able to provide the increases in neutron flux and usable wavelength ranges that will allow biological neutron scatering to reach its true potential.

## 7. References

[1]  I.K. Ibel and H.B. Stuhrmann, *J. Mol. Biol.* **93** (1975) 255

[2]  P.B. Moore, *J. Appl. Crystallogr.* **14** (1981) 237

[3]  G.A. Olah, D.M. Gray, C.W. Gray, D.L. Kergil, T.R. Sosnick, B.L. Mark, M.R. Vaughan and J. Trewhella, *J. Mol. Biol.* **249** (1995) 576

[4]  Y.S. Babu, C.E. Bugg and W.J. Cook, *J. Mol. Biol.* **204** (1988) 191

[5]  J. Trewhella, *Cell Calcium* **13** (1992) 377

[6]  M. Ikura, G.M. Clore, A.M. Gronenborn, G. Zhu, C.B. Klee and A. Bax, *Science* **256,** (1992) 632

[7]  W.E. Meador, A.R. Means and F.A. Quiocho, *Science* **257** (1992) 1251

[8]  D.B. Heidorn, P.A. Seeger, S.E. Rokop, D.K. Blumenthal, A.R. Measn, H. Crespi and J. Trewhella, *Biochemistry* **28** (1989) 6757

[9]  J. Trewhella, D.K. Blumenthal, S.E. Rokop and P.A. Seeger, *Biochemistry* **29** (1990) 9316

[10] G.A. Olah, S.E. Rokop, C.-L.A. Wang, S.L. Blechner and J. Trewhella, *Biochemistry* **33** (1994) 8233.

[11] G.A. Olah and J. Trewhella, *Biochemistry* **33** (1994) 12800

[12] O. Herzberg and M.N.G. James, *J. Mol. Biol.* **203** (1988) 761

[13] K.A. Satyshur, S.T. Rao, D. Pyzalska, W. Drendel, M. Greaser and M. Sundaralingam, *J. Biol. Chem.* **263** (1988) 1628

[14] S.F. Mughabghab, *Neutron Cross Sections: Neutron Resonance Parameters and Thermal Cross Sections*, Vol **1** Parts A and B, Academic Press, London (1984)

[15] J.E. Lynn and P.A. Seeger, *Atomic Data and Nuclear Data Tables* **44** (1990) 191

[16] J. Trewhella, P.A. Seeger, S.E. Rokop, P.D. Palmer, S.J. Henderson and D.E. Hobart, submitted to *Science* (1996)

## DISCLAIMER