

Cryo electron microscopy and electron tomography will play a crucial role in the future of drug development

The completion of the human genome project has focused tremendous interest in determining the structure and function of the tens of thousands of proteins that the genome encodes. Current structural determination techniques are difficult, time-consuming and not generally applicable to all proteins. Cryo Electron Tomography (ET) is, by comparison, relatively straightforward and fast. While it does not achieve atomic scale resolution, it can resolve tertiary and quaternary structure – the level at which much of the critical interaction between proteins occurs. Because it does not require crystallisation, it is applicable to most proteins, including large proteins and protein complexes, flexible proteins and membrane proteins. Because it operates on a single protein molecule, it can provide information that is critical to drug development and unavailable from other techniques, for example, direct observation of interactions among proteins, analysis of the mechanical dynamics and conformation of flexible protein structures, and examination of proteins *in situ*, in the setting in which they naturally occur.

**By Dr Jack Elands
and Dr Ulf Skoglund**

Although the completion of the human genome project has, in theory, provided the amino acid sequence of every protein in the human body, the value of that information will only be realised as we also learn each protein's structure and function. While the sequence is the primary determinant of protein structure, the elucidation of that structure from the sequence alone has proven to be a nearly intractable problem. Indeed, sequence alone may not always be sufficient, as, for instance, when the final structure is determined by post-translational events or interactions with other proteins. There is some hope to be

derived from the knowledge that many proteins share large chunks of similar sequences and structure. For this reason the goal of current efforts to elucidate the human proteome is focused on determining the structure of at least one protein from each 'family', with the expectation that structures of other family members can then be determined by comparisons of sequence alone.

Three-dimensional information of biological systems such as proteins, protein-nucleic acid complexes and macromolecular machines is of critical importance to the understanding of the function. Most of the understanding of molecular events is

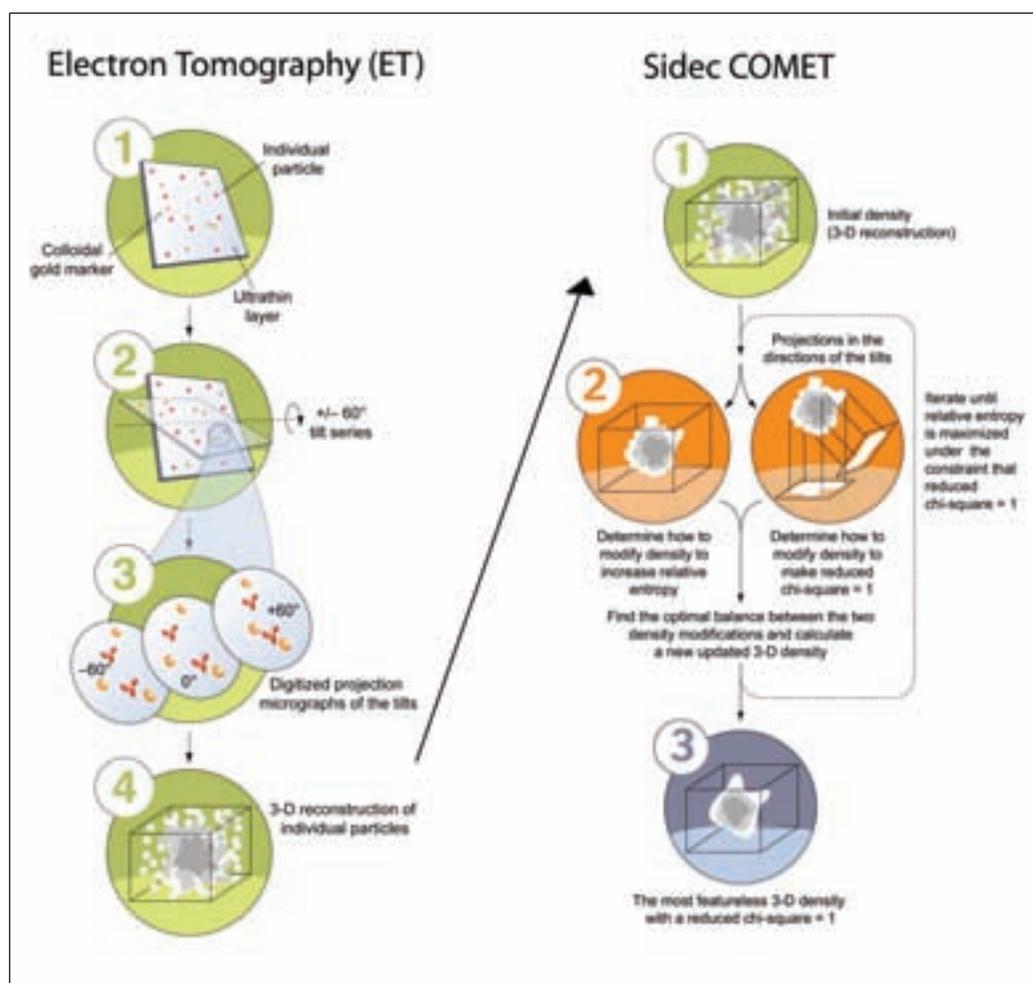


Figure 1 Schematic of SET process. Standard ET uses transmission EM to take a tilt-series of micrographs that are then assembled into 3-D images as a resolution of 5-8nm, revealing cellular features. The Sidec COMET algorithm iteratively refines the image until maximum entropy is obtained to achieve a 3-D reconstruction of the molecules as a resolution of 2-3nm⁴

obtained through various *in vitro* assays, designed to mimic the *in vivo* situation. However, the interaction between macromolecules is dependent on their micro-environment, and *in vitro* assays frequently fail to predict interactions since they are artificial representations of events *in vivo*. The predictable value would be greatly enhanced if the molecular event *in vivo* could be compared to the situation *in vitro*.

The understanding of the actual molecular events and interaction between proteins requires, however, much more than the primary amino acid structure and biochemical assay data. Frequently used methods for structure determination include x-ray diffraction (XRD) and Nuclear Magnetic Resonance (NMR). Both require the expression and purification of relatively large amounts of the subject protein. XRD further requires that the protein be crystallised. Methods for expression and purification are well developed but crystallisation remains impossible for the majority of proteins. Many of the easy to crystallise proteins have

already been analysed so it is not unreasonable to expect declining success in continuing crystallisation efforts. Moreover, crystallisation is certainly not a natural circumstance and may damage or distort native structure. Although NMR does not require crystallisation, it does require intensive calculations that can be equally challenging. Both XRD and NMR become more difficult for larger proteins. Both are slow, typically requiring many weeks to many months for the complete process – expression, purification, crystallisation (XRD), data acquisition and structural analysis.

Though lack of speed and limited applicability may be the most obvious limitation of XRD and NMR, a more fundamental limitation derives from the fact that their analyses necessarily report only an average result based on the summed signal of a large population of individual molecules. They cannot detect the signal of a single molecule nor distinguish the signal of one molecule from another. In practice this prevents their use to analyse molecule to molecule differences in conformation

or interactions between individual molecules. It also precludes their ability to analyse individual molecules *in situ* since they cannot distinguish the target signal from the environmental background.

CryoET can provide conformational analysis with resolution as good as two nanometers, sufficient to define tertiary and quaternary protein structure. Perhaps more significant is CryoET's ability to analyse individual molecules frozen at specific instants in time. This permits the analysis of protein flexibility, direct observations of interactions between proteins and examinations of proteins *in situ*. These capabilities can be thought of as protein imaging. They are of particular value in understanding the interactions between a potential drug and its target, portend an important role for CryoET in drug discovery and development.

Cryo ET

In CryoET transmission electron microscope (TEM) images of a cryogenically frozen specimen are combined to form a three-dimensional model. The TEM works in a way similar to a photographic slide projector: it uses electrons rather than light, the 'slide' is a very thin (less than 100nm) section of the sample material supported on a porous grid,

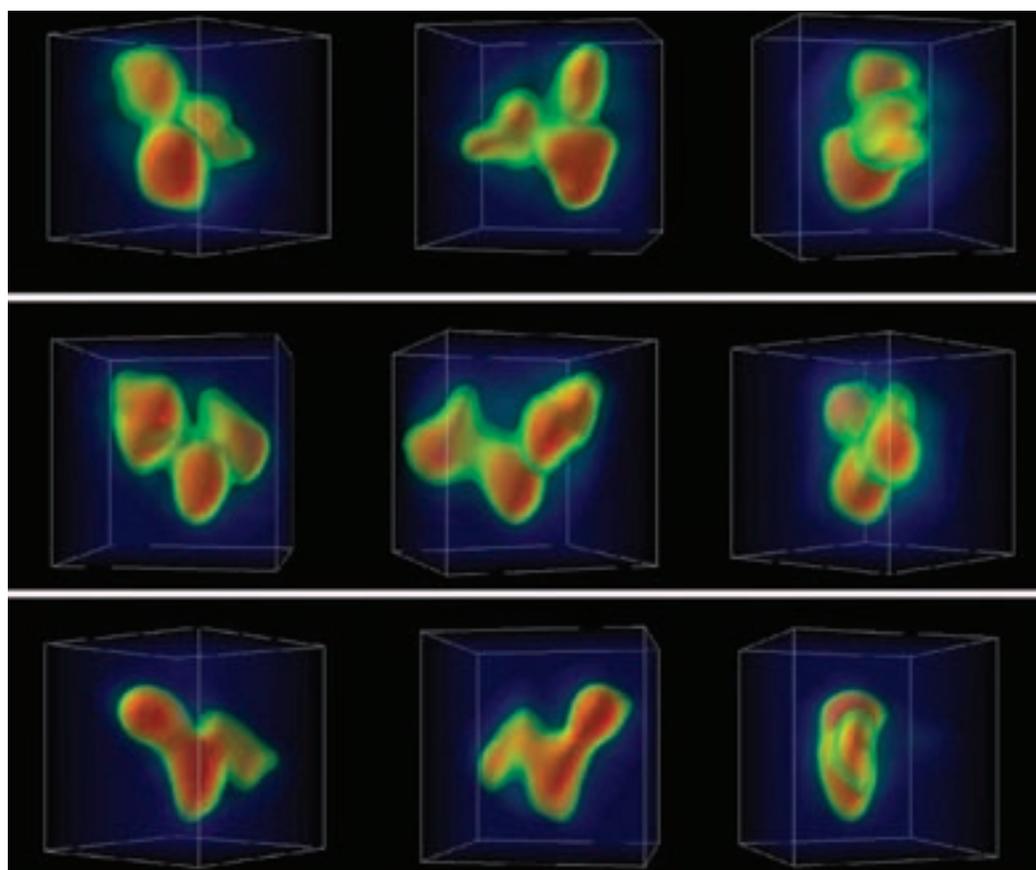
and the image is projected on to a piece of photographic film or an electronic imaging detector. The sample, in this application cells, or a solution of proteins is vitrified by plunging it into liquid ethane at liquid nitrogen temperatures. Cooling is so rapid that water molecules do not have time to crystallise, thus avoiding the damage that would result from normal freezing.

The image produced by the TEM is a two-dimensional projection (on to the plane of the image detector) of the three-dimensional structure of the protein. In electron tomography (ET) the sample is rotated incrementally as a 'tilt-series' of images is acquired, each from a slightly different perspective. The series of projections is then reconstructed into a three-dimensional model of the sample. The computational algorithms for this reconstruction have been highly refined in recent years, primarily through their use in medical macro-imaging techniques such as x-ray, CAT scanning and MRI. Although TEM itself is capable of resolution better than 0.1nm, the three-dimensional reconstructions generated by ET have resolutions of about 2nm.

Key to the recent success of ET for protein structural analysis has been the development of

Figure 2

Conformations of three individual IgG molecules determined by CryoET. Each molecule (a-c) is displayed from three different perspectives. Observation of a large number of such tomograms permits a statistical analysis of shape distribution and the derivation of a potential energy function (Bongini et al¹)



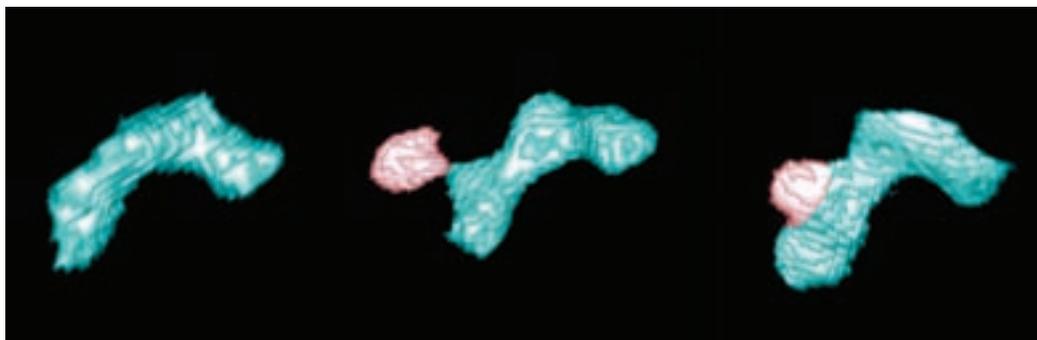


Figure 3
ET was used to study the interaction of ScFv, a single chain antibody, with human complement factor C5. On the left is a reconstruction of human complement factor C5. The middle image shows ScFv (red) bound to human complement factor C5 (blue). The image on the right shows both proteins again but the interacting areas are larger (Savage²)

computational techniques to improve the signal-to-noise ratio and resolution of the model. Proteins are relatively delicate structures that can be easily destroyed by the 200-300 KeV electrons of the TEM beam. To avoid damage, exposure must be limited to a few hundred electrons per square nanometer. Under these conditions the signal is very low relative to the noise and simple averaging cannot offer sufficient contrast enhancement without exceeding the permitted electron dosage. COMET (Siedec Technologies AB, Stockholm, Sweden) is an imaging algorithm designed to enhance contrast and resolution in electron tomography. Starting with an initial three-dimensional model, the iterative mathematical procedure maximises the entropy of the model under the constraint of a χ^2 fitting parameter, essentially producing the most featureless model that still fits the observed projection data. The technique can be applied to ET data from thin sections that have been vitrified or 'fixed' by other preparation techniques, and is, in fact, generally applicable to tomographic data of all types.

Analysis of flexible proteins

Proteins are typically not rigid. In fact their flexibility is probably essential to many of their functions. For example, flexibility in antibodies promotes their ability to connect antigens to immunological effectors. The antigen binding sites are free to twist and turn in order to match up with a wide variety of antigens and link them to a relatively small number of effectors. Understanding flexibility and the dynamics of protein shape changes is essential to understanding interactions between proteins. Unlike XRD and NMR, all data in a CryoET analysis comes from the same molecule. By comparing the models from a population of flexible molecules, each in a slightly different shape, it is possible to analyse the statistical distribution of shapes and estimate the potential energy of the molecule.

In a recent study L. Bongini et al¹ examined immunoglobulin G (IgG), a Y-shaped molecule composed of two fragment antigen-binding arms (Fab arms) connected by flexible hinges to an effector stem (Figure 2). The investigators developed a model for the potential energy of IgG. They noted that the average angle between the Fab arms (110°) differs significantly from the value determined by XRD (172°) and suggested that the difference is due to the packing environment of the crystal. It is reasonable to suggest that such distortion may be generally present in XRD structures and that the more flexible the protein, the larger and more likely the errors.

Observing interactions

Closely related to its ability to study molecular dynamics, is CryoET's ability to investigate protein interactions. Figure 3 demonstrates the technique's ability to visualise the binding site and binding dynamics of an antibody with its target antigen², in this case ScFv, a small (30kDa) single chain antibody that binds to human complement factor C5. The investigators were able to visualise the epitope and confirm that the shape of the bound complex agreed with their modelling results. In a drug development programme, protein imaging studies such as this one could allow the selection of the most appropriate antibody from an available library based on the antibody's ability to bind to a particular antigen.

In situ analysis

Perhaps the most important capability of CryoET is its ability to visualise protein conformations *in situ*, in their native biological context. We are used to seeing neat, clean, textbook diagrams of isolated proteins and it is easy to forget that the environment in which they operate is crowded not only with other proteins, but also with cell organelles, membranes and other organic and inorganic molecules. Understanding protein function is largely a

matter of understanding how proteins relate to and interact with these other cellular components. Although *in situ* analysis lets us observe these interactions directly, the analysis is also made more complicated by the fact that now we must distinguish the molecule of interest from everything else. The easiest way to do so is the use of immunolabelling techniques. An antibody specific to the protein is attached to a gold particle a few nanometers in diameter. Because the gold has a high density relative to the biological material, it is easy to see in the TEM images. Gold particles are also used as reference markers for proper alignment of the series of tilted images allowing for a more accurate reconstruction.

Many of the most interesting proteins are membrane proteins. Because they are ‘designed’ to exist embedded in the hydrophobic environment of the cell membrane, it is unlikely that they will ever be crystallised for x-ray analysis in their native form. They also tend to be relatively large for NMR.

Nephrin is a membrane protein known to play a critical role in the filtration of water and waste from the blood in the kidney. Patients with nephrin defects can suffer from proteinuria as protein leaks through the defective filtration network and into the urine. Filtration occurs in the glomerulus of the kidney and the filter itself is three layered: the capillary endothelium, the glomerular basement membrane, and a layer of special cells called podocytes. The podocytes cover the surface of the capillary with specialised interdigitated foot processes. Filtration slits occur between the foot processes and are bridged by slit diaphragms.

Nephrin has long been associated with the slit diaphragms. It is an elongated globular protein comprising an intracellular domain within the podocyte, a transmembrane domain crossing the podocyte cellular membrane, and an extracellular domain that actually constitutes the filtration network. Using electron tomography, Wartiovaara et al³ were able to visualise the interlocking filtration network formed by the extracellular domains of nephrin molecules (Figure 4).

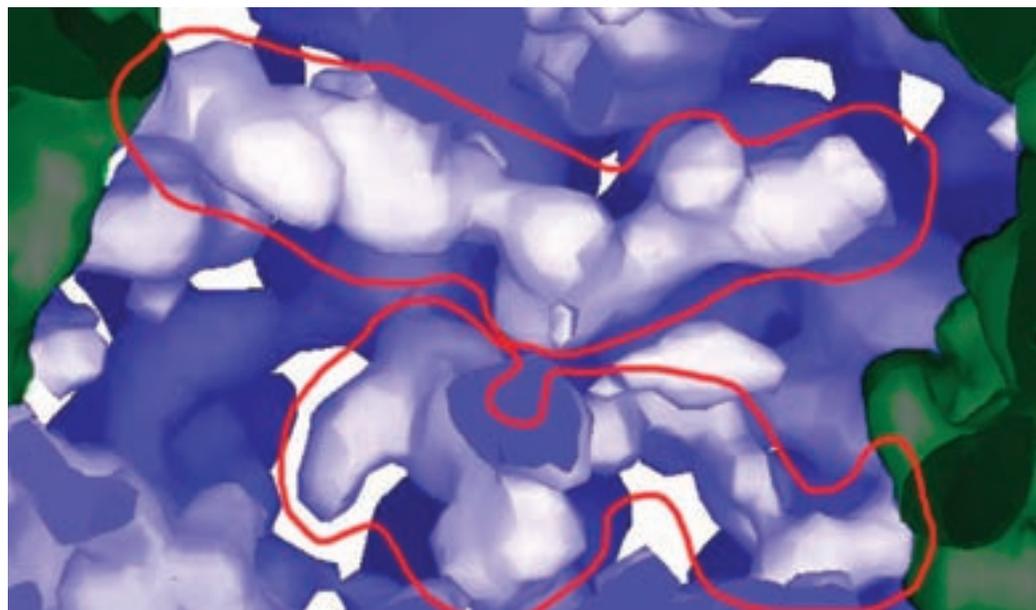
Banyay et al⁴ used electron tomography to analyse the protein conformation of an ion channel. The study is noteworthy not only because the subject proteins were membrane proteins analysed *in situ*, but also because they were large complexes of multiple subunits. The investigators were able to determine the conformation of the fully assembled channel and demonstrate the aggregation of protein subunits into the final channel assembly as a tetramer. They conclude: “Understanding protein stoichiometry and subunit composition *in situ* is particularly important in the target definition process, being crucial information in setting up a relevant screening assay. The results of this study demonstrate that ET is a valuable tool to visualise protein conformation of *in situ* samples, thus affording crucial information about ion channels that can enable improved assay screening systems in drug development.”

Conclusion

Historically most drugs were discovered serendipitously – some agent was observed to have a desirable therapeutic effect and it was administered

Figure 4

Tomogram of human filtration slit, front view. Globular cross strands about 35nm in length (outlined in red) extend from podocyte membranes (Wartiovaara³)



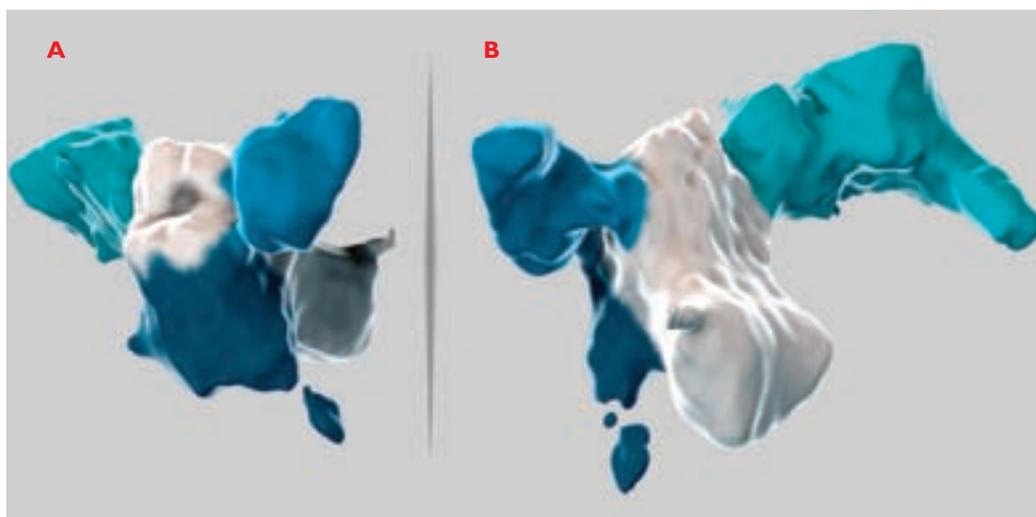


Figure 5: Ion channel expressed on the plasma membrane of RIN cells. Analysis of the size and structure of the macromolecular complex indicates a tetrameric structure spanning approximately 15-17nm in height and 9-10nm in width. The channel is pictured in white with extracellular components pointing upwards and intracellular components downwards in (A). This view clearly depicts the extracellular side of the ion channel that resembles the crown of a tooth with a central pore marked in gray. (B) is a clockwise rotation of the ion channel. (Images generated using 4D Cinema, Maxon, Friedrichsdorf, Germany) (Banyay⁴)

without any real understanding of its operating mechanism. Only in the last 50 years have we come to understand the mechanisms in some detail, and only in the last decade have we understood many at the molecular level. We are now faced with a veritable explosion of molecular detail that will dramatically affect the way we search for new drugs. The inherent economic inefficiencies of brute force and serendipitous discovery models will lead inevitably to their displacement by rational drug design. But success in rational design requires much more than simple knowledge of static protein structure and conformation. It requires also a detailed understanding of intra- and intermolecular dynamics as proteins interact with other proteins and with the myriad other components of the cellular environment. CryoEM with its unique ability to provide time-specific snapshots – images – of single molecule protein conformations within the native cellular milieu will play a critical enabling role in the ascendancy of rational drug design. **DDW**

Dr Jack Elands obtained his PhD in medical pharmacology from the University of Utrecht (Netherlands) and the CNRS-INSERM Institute for Pharmacology and Endocrinology in Montpellier (France). At Marion Merrell Dow (now Sanofi-Aventis) he helped the company in its transit to embrace a more industrialised discovery research approach. He was also responsible for

several candidates in late research and early clinical research. Dr Elands has held strategic marketing and business development positions at Zymark (now Caliper LS), Packard Bioscience (now PerkinElmer) and IDBS. He is currently a company director at FEI and responsible for life science business development.

Dr Ulf Skoglund is Principal Scientist and Founder of Sidec Technologies AB and also Professor of Molecular Genetics at the Department of Cell and Molecular Biology, Karolinska Institute. Dr Skoglund received his PhD in structural chemistry from Stockholm University (Sweden). His research group has developed the electron tomography techniques over the last two decades and has maintained a leading position in the field.

References

- 1 Bongini, L et al. Freezing immunoglobulins to see them move. PNAS, April 27, 2004, vol. 101, no. 17, pp. 6466-6471.
- 2 Savage, C. 3-D imaging of Bio-molecules, Innovations in Pharmaceutical Technology, January 2003, pp. 56-60.
- 3 Wartiovaara, J et al. Nephin strands contribute to a porous slit diaphragm scaffold as revealed by electron tomography. J Clin Inv, 114:10 November 2004
- 4 Banyay et al, M. Three-dimensional imaging of *in situ* specimens with low-dose tomography to analyze protein conformation, Assay and Drug Development Technologies, 2:5, 2004.