

Scripps Cryo Course, November 2005

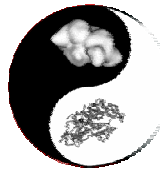
## COMBINING ELECTRON MICROSCOPY WITH ATOMIC MODELS

Niels Volkmann  
The Burnham Institute for Medical Research  
10901 North Torrey Pines Road.  
La Jolla, California, USA  
niels@burnham.org

Electron Microscopy can give structural information on many complex systems but is most often limited to non-atomic resolution (usually between 10-30Å)

Techniques for determination of atomic structures are limited by size or crystallinity requirements

We can gain atomic-level information on large complexes by docking atomic models of components into lower-resolution reconstructions from electron microscopy

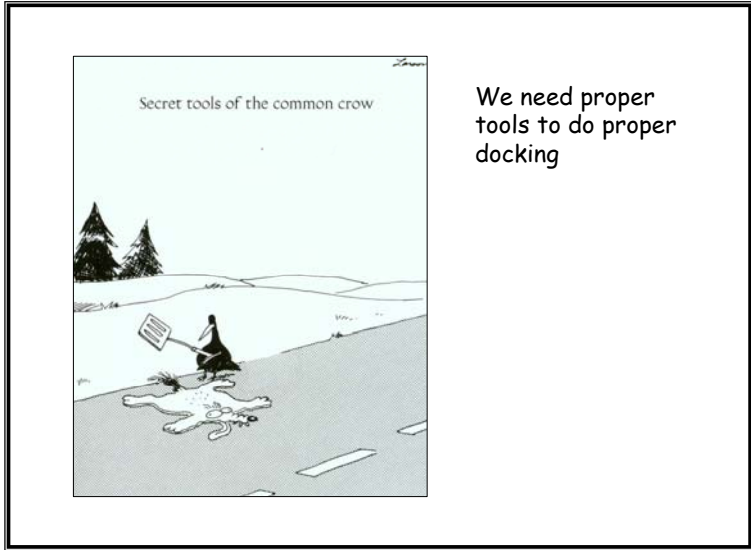


All we need to do is to stick an atomic model into the EM density, wiggle and deform it to fit the density better and then we have atomic resolution information, right?

**No, wrong!**

We try to find the correct positions of highly localized atoms within a relatively featureless density (Atoms into Blobs).

Finding a “perfect fit” is the easy part, figuring out if the fit is meaningful, that is the hard part.



We need proper tools to do proper docking

**Part I:**  
**Sticking in the model**

## **Approaches to docking**

### **Manual docking**

- Immediate visual feedback
- Heavy human intervention
- High level of subjectivity
- Prone to biasing
- Dependent on contour level

## **Approaches to docking**

### **Landmark-based docking**

- Reduced representation, therefore fast
- Moderate human intervention
- Loss of data
- Error of docking position hard to assess
- Needs one-to-one correspondence of map and model

## **Approaches to docking**

### Density-based docking

Does not need one-to-one correspondence of map and model

Can potentially handle modular docking

Little human intervention

Density data fully explored

Independent of contour level

Relative expensive calculation, can be slow

## **Approaches to docking**

### Surface-based docking

Can potentially handle modular docking if modules have distinct surface features

Little dependency on internal features

Density data not fully explored

Corresponds to high-pass filtering, therefore potentially error prone

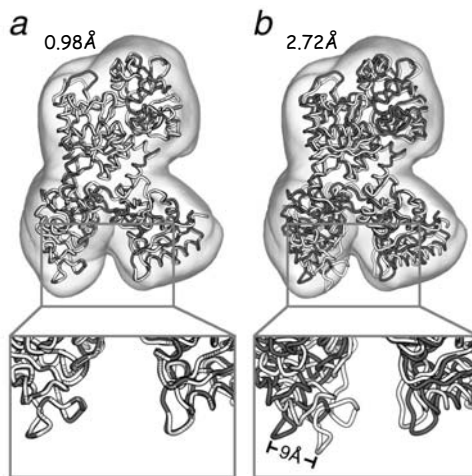
Expensive calculation, can be slow

## Approaches to docking

### Local refinement, flexible docking

- Can potentially account for local variations
- Can use additional information (stereo chemistry, normal modes)
- Reduces observable to parameter ratio
- Serious danger of over-fitting

## MODULAR VERSUS FLEXIBLE



Wriggers & Birmanns, JSB 2001  
Volkmann & Hanein, Meth Enzym, 2004

## **Approaches to docking**

### Use of other filters

*Masking* can enhance performance for relatively noise-free maps

Corresponds somewhat to high-pass filtering and is thus susceptible to high frequency noise

*Core weighting* can improve performance for multimers with certain shape characteristics

Filters usually slow down calculation

## **Approaches to docking**

Compare atomic positions directly with density

Convert atomic model to density, then compare

Convert atomic model and density to something else, then compare

**Part II:  
Evaluating the quality  
of the fit**

**Similarity Measures**

Measure the fit between atomic model and reconstruction

Maximum Likelihood gives the best possible unbiased estimate (Neyman and Pearson)

Many of the more common similarity measures are some approximation to maximum likelihood

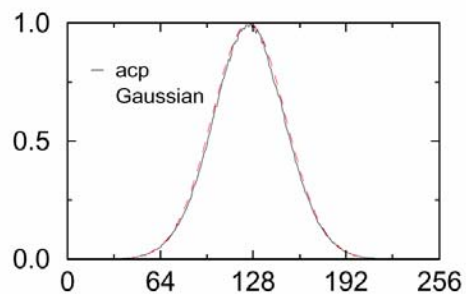
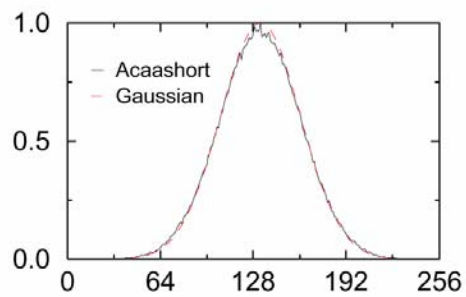


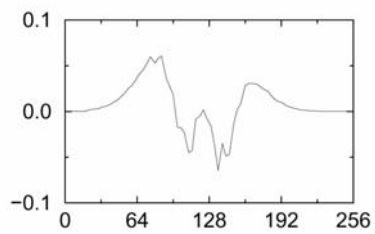
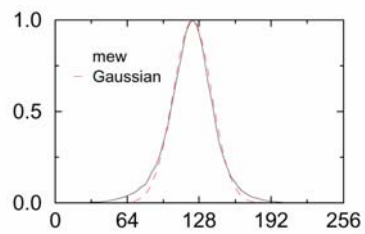
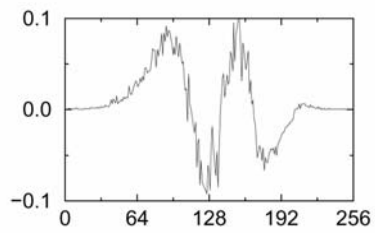
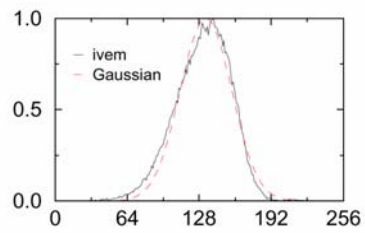
## Correlation Coefficient

$$CC = \frac{\sum (a_i - \bar{a})(e_i - \bar{e})}{\sqrt{\sum (a_i - \bar{a})^2 \sum (e_i - \bar{e})^2}}$$

CC is a Maximum Likelihood measure if noise is uniform Gaussian and  $a$  and  $e$  are independent, related by an affine relationship (i.e.  $e = ca + d$ )

*Cross-correlation* is not a maximum likelihood measure; requires an identity relationship between  $a$  and  $e$  (i.e.  $e = a$ )





Once we decided on a particular docking approach and a particular similarity measure, how do we proceed?



**Part III:**  
**Finding confidence intervals**

**Numerical Recipes**

To be genuinely useful a docking procedure should provide:

- (i) accurate, globally best parameters
- (ii) error estimates on these parameters
- (iii) a statistical measure for goodness-of-fit

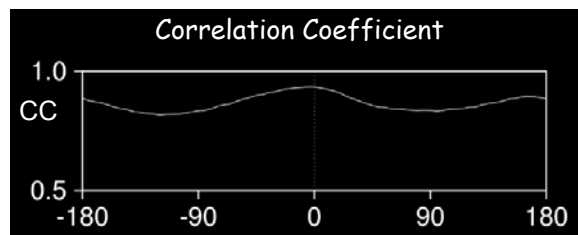
“Unfortunately many practitioners of parameter estimation never proceed beyond item (i)! They deem a fit acceptable if a graph of data and model ‘looks good’. This approach is known as chi-by-eye.”

Addressing (ii) and (iii):

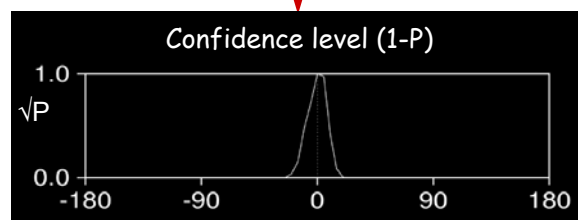
### **Solution Sets**

Define confidence level at which solutions are still equivalent to the fit with the globally highest score of the similarity measure.

Statistically, all the solutions within this set satisfy the data equally well and have equal probability, at the chosen confidence level, to be the 'true' solution.



Fisher's z transform

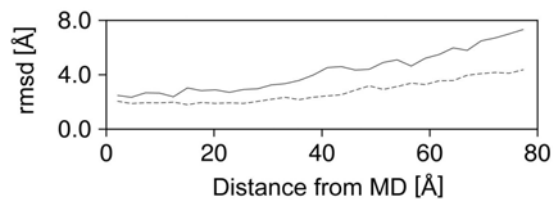
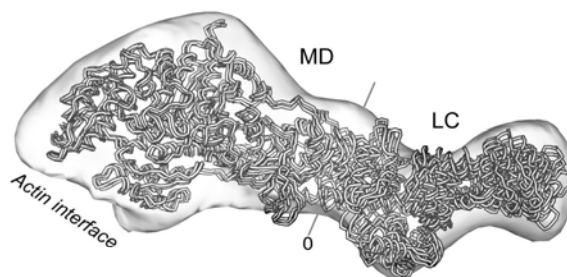


## Solution Sets

The sets can be used to calculate parameters of interest as a property of the set, for example the

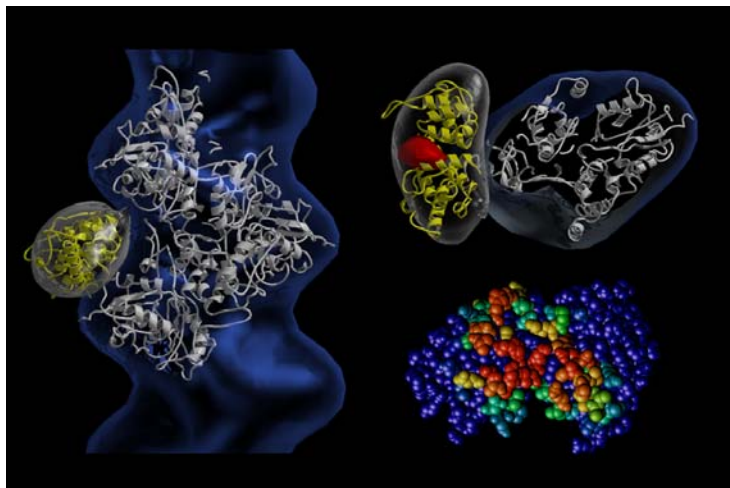
Coordinate error can be estimated by the rmsd value of the atom position within the whole set

## CONFORMATIONAL VARIATION



## **Solution Sets**

Interaction probabilities, the probability that certain residues take part in interactions, can be estimated by integrating over the solution sets

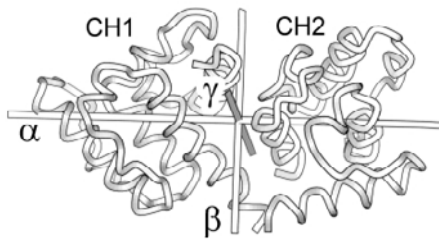


## Solution Sets

Degeneracies can be easily detected by analyzing solution sets

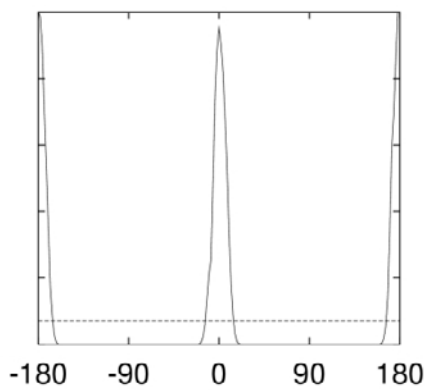


## DETECTION OF DEGENERACIES: LOW-RESOLUTION PSEUDO SYMMETRY

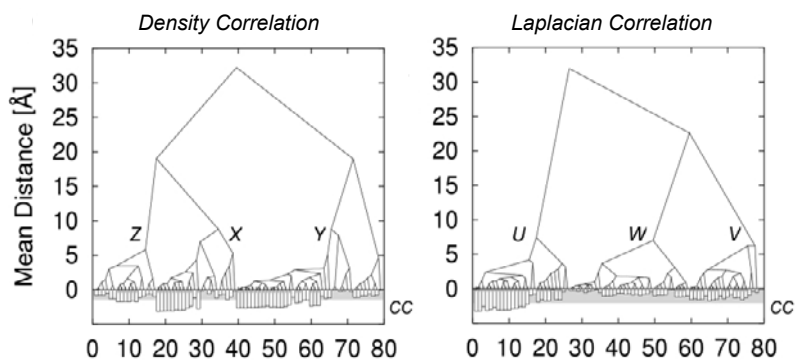


*Fimbrin*

### DETECTION OF DEGENERACIES: LOW-RESOLUTION PSEUDO SYMMETRY

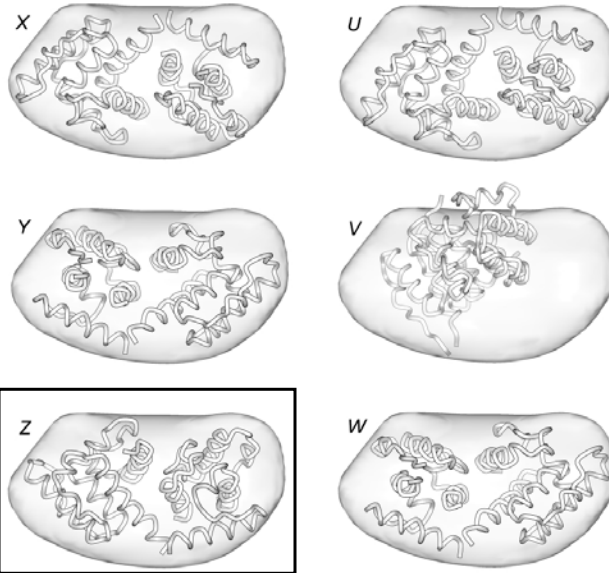


### DETECTION OF DEGENERACIES, CLUSTER ANALYSIS





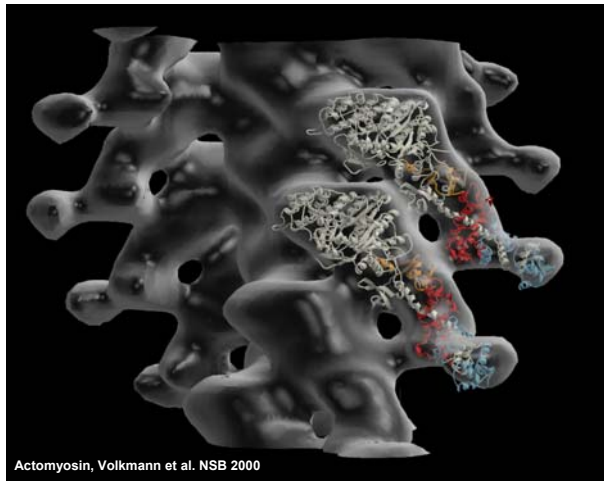
## DENSITY CORRELATION VERSUS LAPLACIAN FILTER



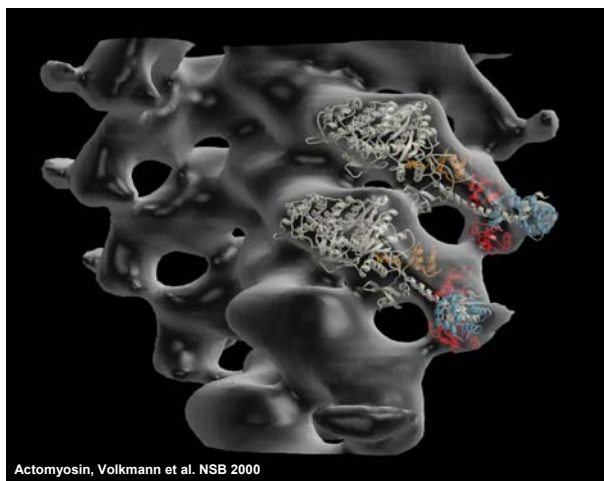
### **Application Example**

Helical reconstruction of actin-bound smooth muscle myosin

- Step 1: isolate myosin contribution from reconstruction
- Step 2: define modules
- Step 3: dock largest module
- Step 4: subtract contribution of docked module



Actomyosin, Volkmann et al. NSB 2000



Actomyosin, Volkmann et al. NSB 2000

## Acknowledgements

### *Volkman Lab:*

Xiao-Ping Xu  
Christopher Page  
Martin Fleming  
Susanta Mukhopadhyay  
Zahir Basrai  
Jonathon Sexton  
Emilie Perlade

### *Collaborators:*

Dorit Hanein  
Susan Lowey  
Kathy Trybus  
Andreas Hoenger  
Eckhard Mandelkow

NIH funded postdoctoral positions available...

Contact: [niels@burnham.org](mailto:niels@burnham.org)



The Burnham Institute for Medical Research- La Jolla, California



72°