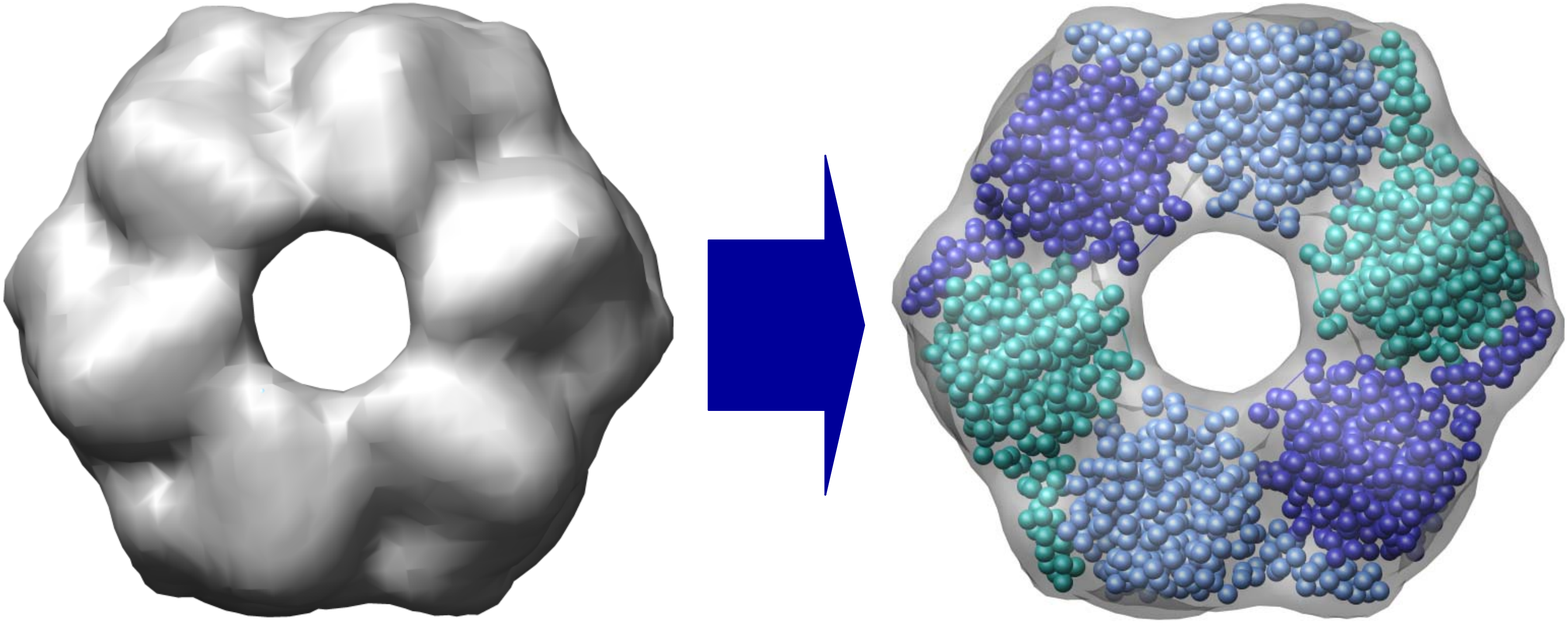
deus ex biomachina

biomachina.org

# Applications of Fourier Theory: Template Matching

Jochen Heyd
School of Health Information Sciences &
Institute of Molecular Medicine
University of Texas – Houston

# The Goal



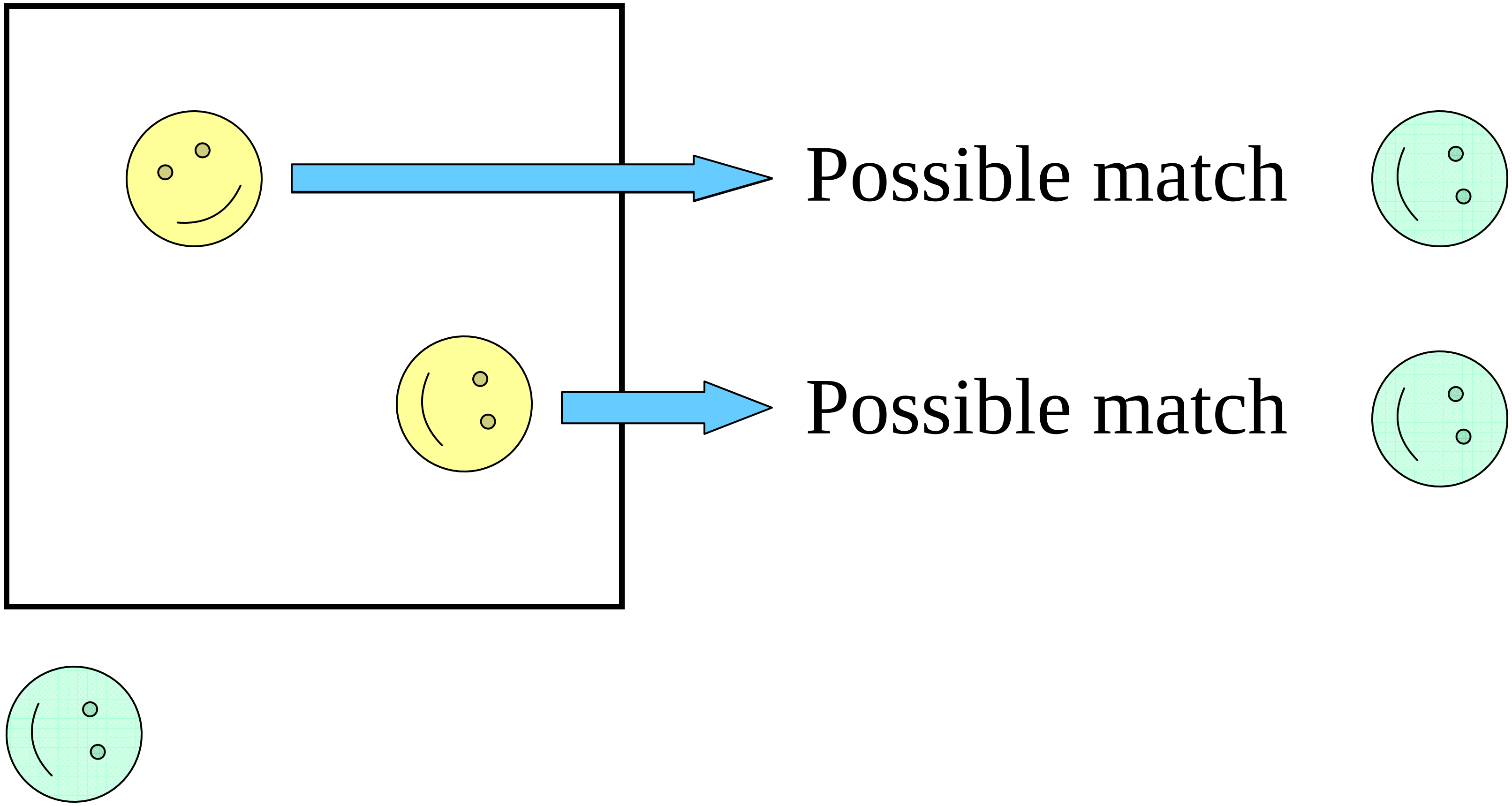
EM Density

Atomic-level Description

# Contents

- Template convolution
- Fast translational matching
- The correlation landscape and ways to improve it
- Assessment using simulated data
- Application to experimental microtubule data

# Template “Convolution”: An Example

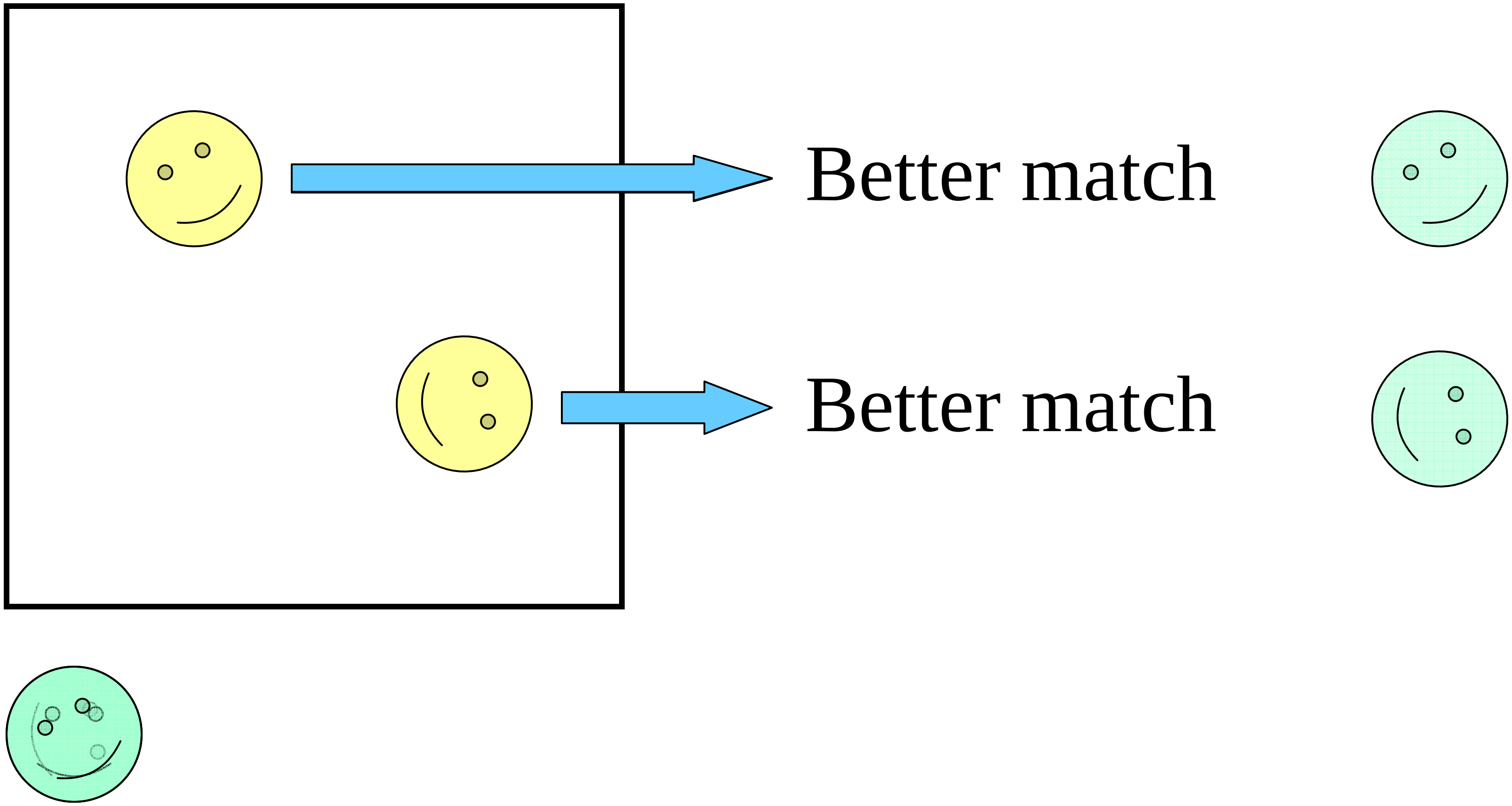


Possible match

Possible match

Scoring function: cross-correlation

# Template “Convolution”: An Example



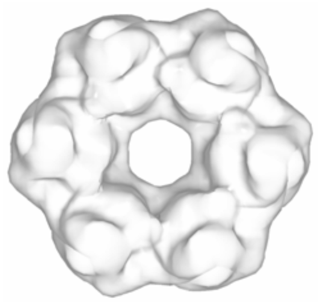
Better match

Better match

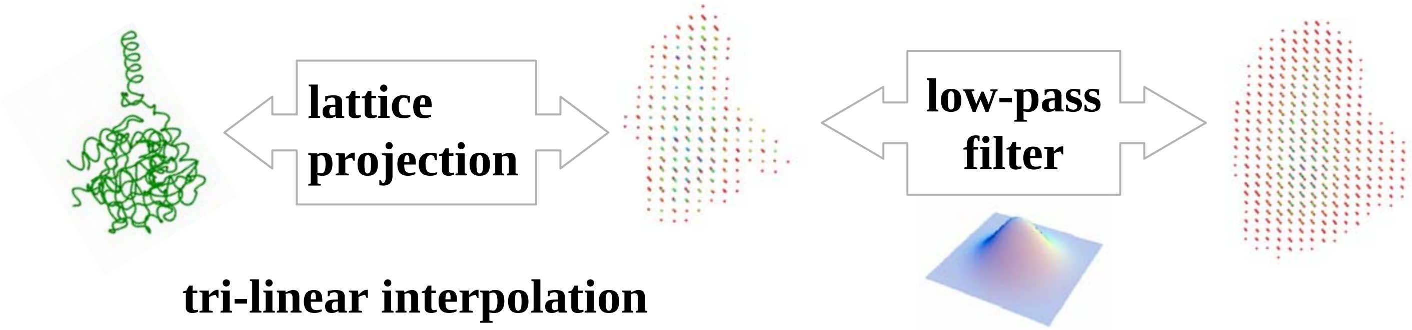
Scoring function: cross-correlation

# Template "Convolution"

$\rho_{\text{em}}(\mathbf{r})$ target density on lattice



$\rho_{\text{atomic}}(\mathbf{rR})$ rotated probe molecule density projected to the lattice:



**lattice projection**

**low-pass filter**

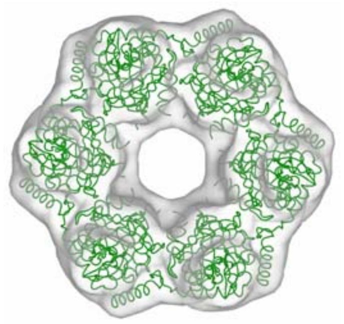
$$g \otimes \rho_{\text{atomic}}(\mathbf{rR})$$

$$\equiv \rho_{\text{calc}}(\mathbf{r})$$

**tri-linear interpolation**

***g* = Gaussian kernel**

$$\boxed{C(\mathrm{T}) = \int \rho_{\text{em}}(\mathrm{r}) \cdot \rho_{\text{calc}}(\mathrm{r}+\mathrm{T})\, d^3\mathrm{r}} \Longrightarrow$$



Fitting criterion: e.g. linear cross-correlation,
evaluate for every rotation **R** and translation **T**

P. Chacón

# Computational Cost

- Three translational degrees of freedom

    N possible locations

- Three rotational degrees of freedom

    M possible orientations

- Cost for each cross-correlation calculation

    N (number of voxels)

Total cost: $N * M * N = M * N^2$

# FTM (Fast Translational Matching)

The expression for the cross-correlation is

$$C(\mathrm{T}) = \int \rho_{\mathrm{em}}(\mathrm{r}) \cdot \rho_{\mathrm{calc}}(\mathrm{r}+\mathrm{T})\, d^{3}\mathrm{r}$$

Using the Fourier Convolution Theorem, we get

$$\boxed{C(\mathrm{T}) = \mathrm{F}^{-1}\left[\ \mathrm{F}(\rho_{em})^{*} \cdot \mathrm{F}(\rho_{calc})\ \right]}$$

Needs to be calculated only ONCE

This yields ALL possible translations in one step!

# FTM (Fast Translational Matching)

The expression for the cross-correlation is

$$C(\mathrm{T}) = \int \rho_{\mathrm{em}}(\mathrm{r}) \cdot \rho_{\mathrm{calc}}(\mathrm{r} + \mathrm{T})\, d^{3}\mathrm{r}$$

Using the Fourier Convolution Theorem, we get

$$\boxed{C(\mathrm{T}) = \mathrm{F}^{-1}\left[\ \mathrm{F}(\rho_{em})^{*} \cdot \mathrm{F}(\rho_{calc})\ \right]}$$

Need to be calculated for every orientation

This yields ALL possible translations in one step!

# Computational Cost

- ~~Three translational degrees of freedom~~

  ~~N possible locations~~

- Three rotational degrees of freedom

  M possible orientations

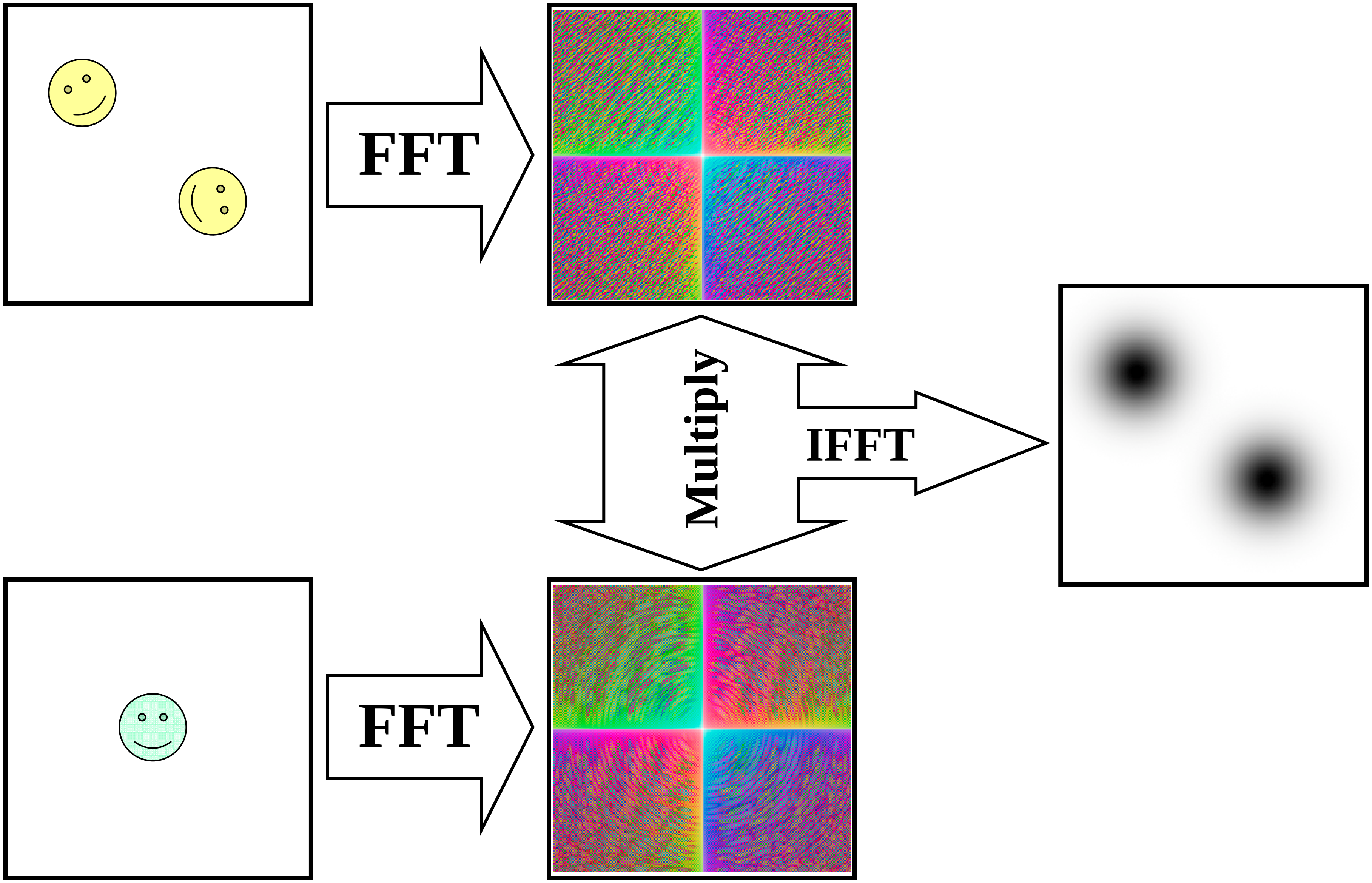
- Cost for each cross-correlation calculation

  ~~N~~

→ Total cost: M * 2 N log(N)

For a $50^3$ map this results in a speedup of 4 orders of magnitude!

# FTM: An Example



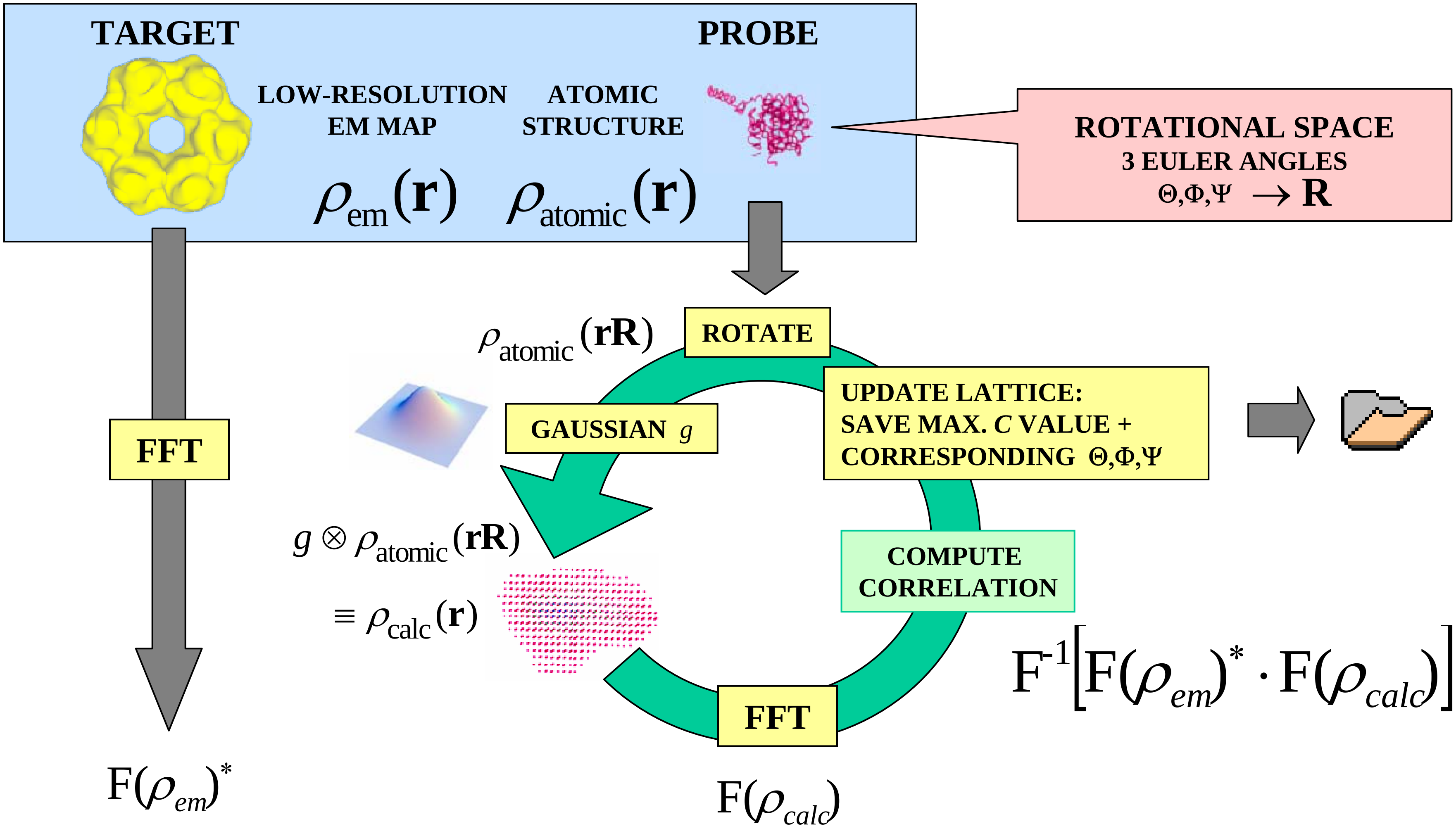
FFT

Multiply

IFFT

FFT

# 6D Search with FTM



**TARGET**

**LOW-RESOLUTION EM MAP**

$\rho_{em}(\mathbf{r})$

**PROBE**

**ATOMIC STRUCTURE**

$\rho_{atomic}(\mathbf{r})$

**ROTATIONAL SPACE**
**3 EULER ANGLES**
$\Theta,\Phi,\Psi \rightarrow \mathbf{R}$

**FFT**

$F(\rho_{em})^*$

**ROTATE**

$\rho_{atomic}(\mathbf{rR})$

**GAUSSIAN** $g$

$g \otimes \rho_{atomic}(\mathbf{rR})$

$\equiv \rho_{calc}(\mathbf{r})$

**FFT**

$F(\rho_{calc})$

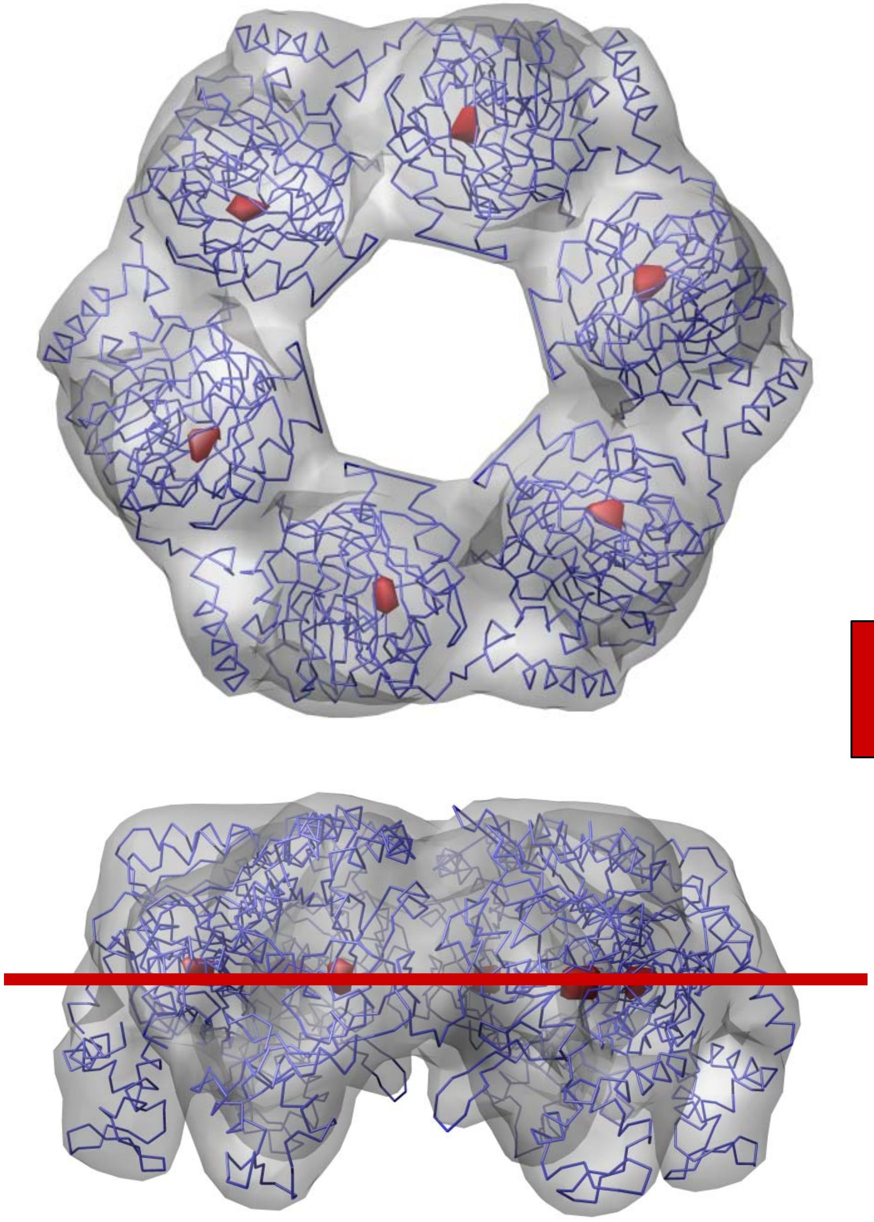
**COMPUTE CORRELATION**

$$F^{-1}\left[F(\rho_{em})^* \cdot F(\rho_{calc})\right]$$

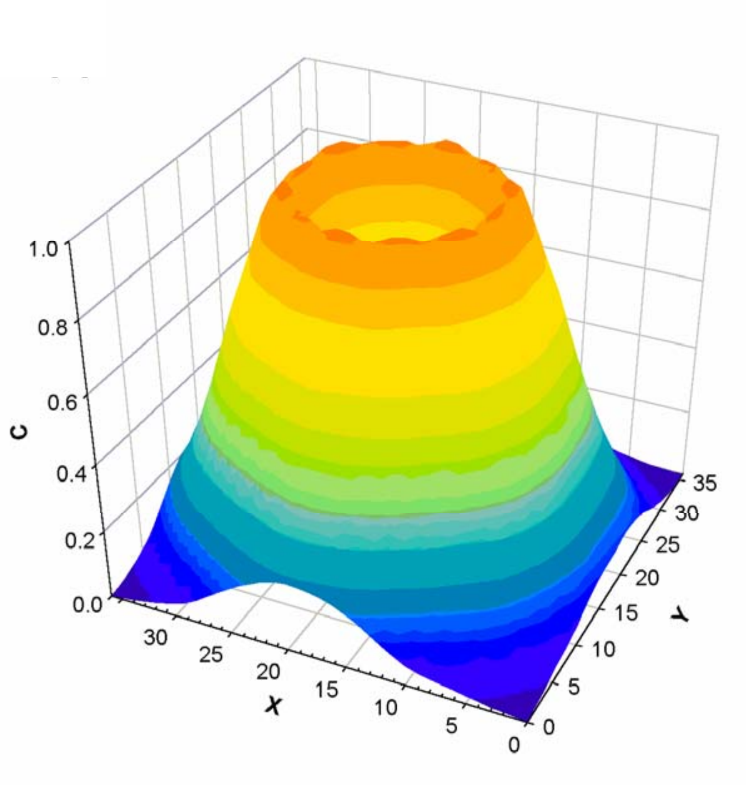
**UPDATE LATTICE:**
**SAVE MAX. *C* VALUE +**
**CORRESPONDING $\Theta,\Phi,\Psi$**

P. Chacón

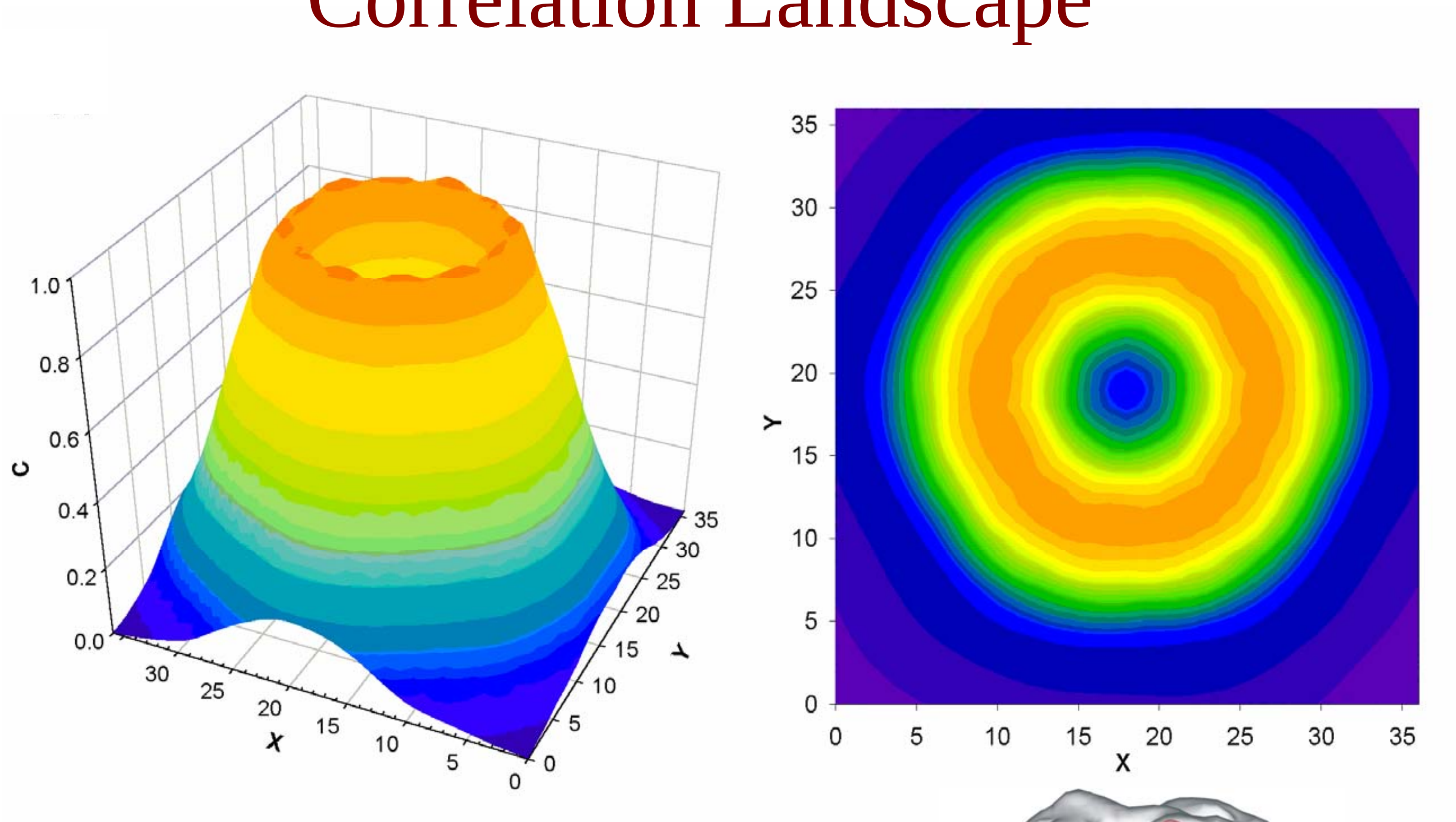
# Correlation Landscape



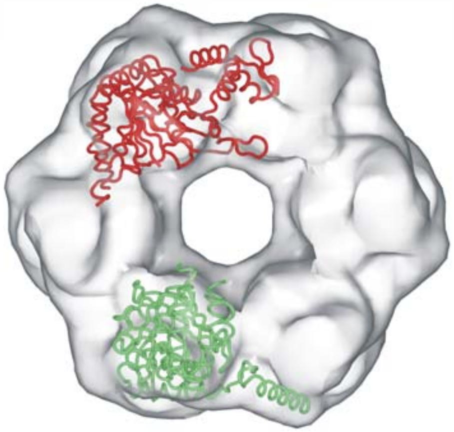
Translation Function



# Correlation Landscape



**With density cross-correlation we cannot distinguish between correct and spurious fit**



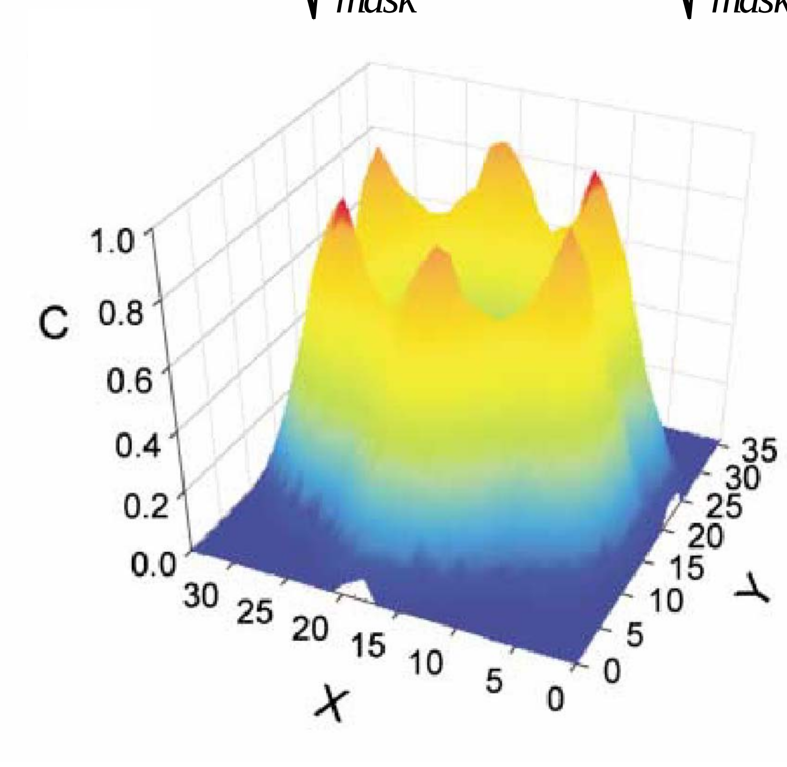
P. Chacón

# Density Masking

Renormalize (mask) the correlation locally:

$$C(\mathbf{T}) = \frac{\int_{mask} \rho_{\mathrm{em}}(\mathbf{r}) \times \rho_{\mathrm{calc}}(\mathbf{r}+\mathbf{T})\ \mathrm{d}^3\mathrm{r}}{\sqrt{\int_{mask} \rho_{\mathrm{em}}^2(\mathbf{r})\mathrm{d}^3\mathrm{r}}\sqrt{\int_{mask} \rho_{\mathrm{calc}}^2(\mathbf{r})\ \mathrm{d}^3\mathrm{r}}}$$

$$\mathrm{mask} \rightarrow \rho_{\mathrm{calc}_{l,m,n}} > 0$$



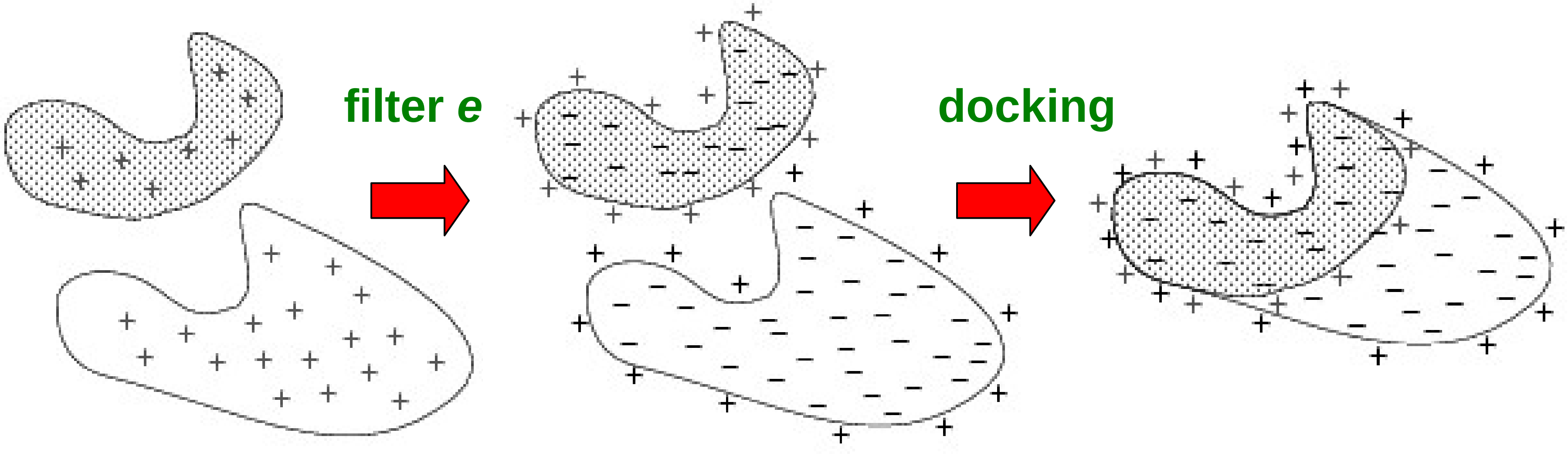
- Extends the reliability of correlation based docking (<15Å)
- Cannot be easily FTT accelerated

DOCKEM, A.M. Roseman

# Density Filtering

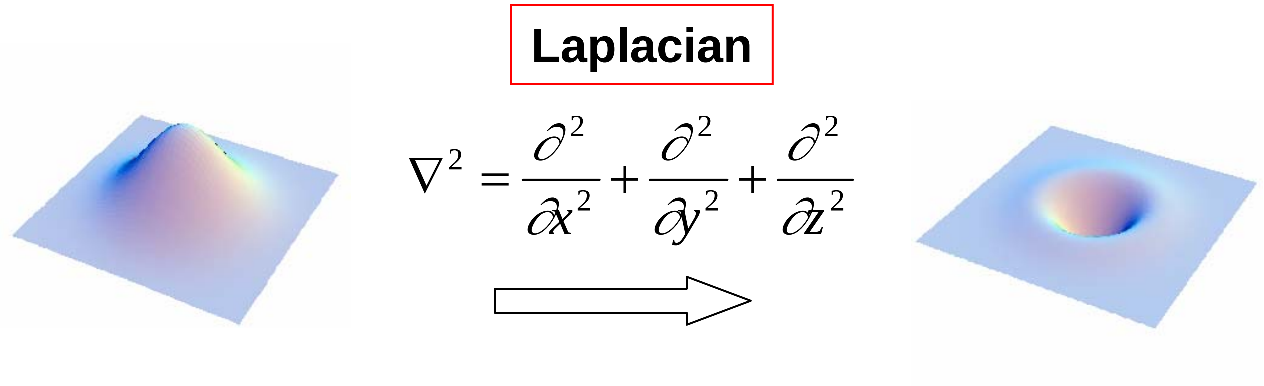
## Adding surface/contour information

A suitable filter would assign negative values to the interior, positive values to the molecular contour. Both volume and contour matches would provide positive contributions to the correlation criterion:



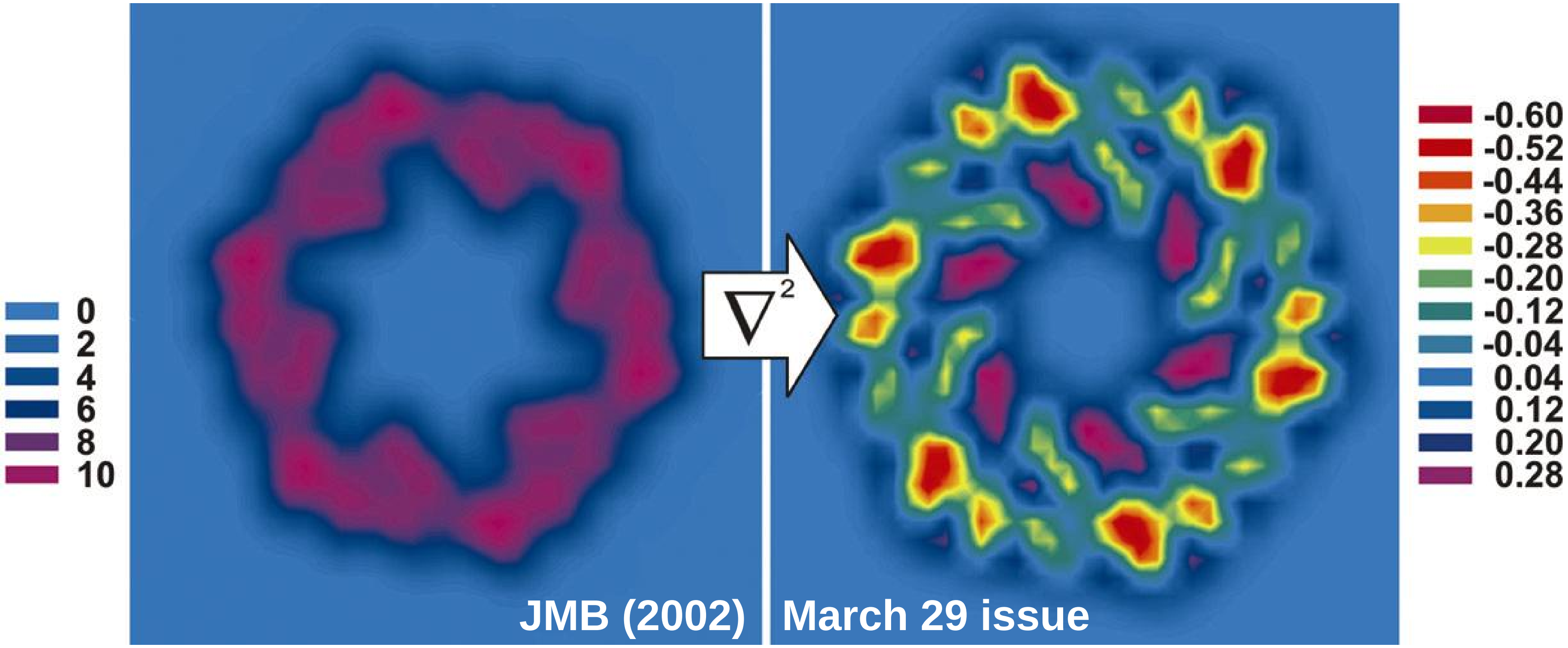
P. Chacón

# Contour Filter



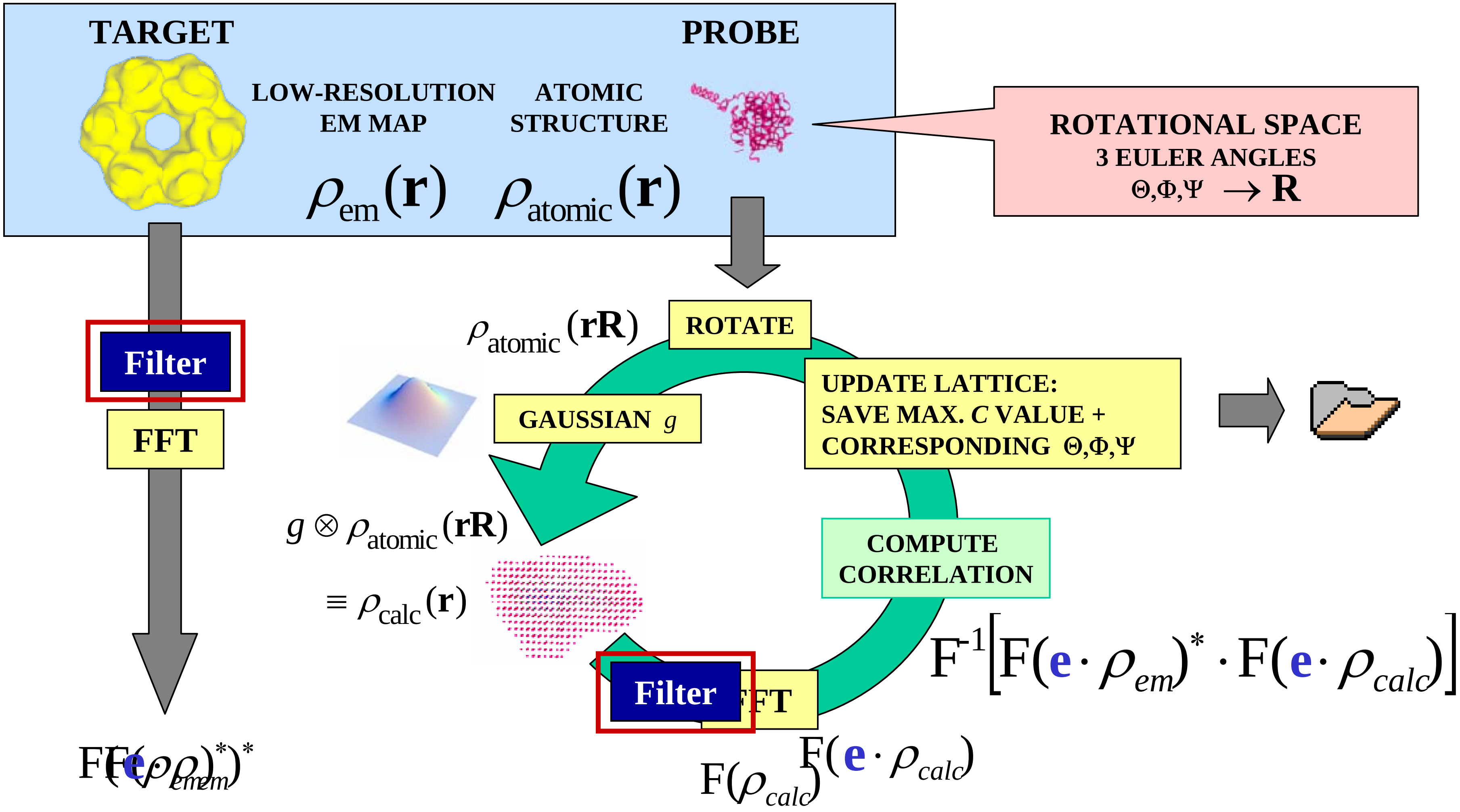
**Laplacian**

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$



P. Chacón

# 6D Search with FTM and Filtering



**TARGET**

LOW-RESOLUTION
EM MAP

$\rho_{em}(\mathbf{r})$

**PROBE**

ATOMIC
STRUCTURE

$\rho_{atomic}(\mathbf{r})$

**ROTATIONAL SPACE**
**3 EULER ANGLES**
$\Theta,\Phi,\Psi \rightarrow \mathbf{R}$

**Filter**

**FFT**

$F(\mathbf{e}\cdot\rho_{em})^*$

$F(\rho_{em})^*$

**ROTATE**

$\rho_{atomic}(\mathbf{rR})$

**GAUSSIAN** *g*

$g \otimes \rho_{atomic}(\mathbf{rR})$

$\equiv \rho_{calc}(\mathbf{r})$

**Filter**

**FFT**

$F(\rho_{calc})$

$F(\mathbf{e}\cdot\rho_{calc})$

**COMPUTE CORRELATION**

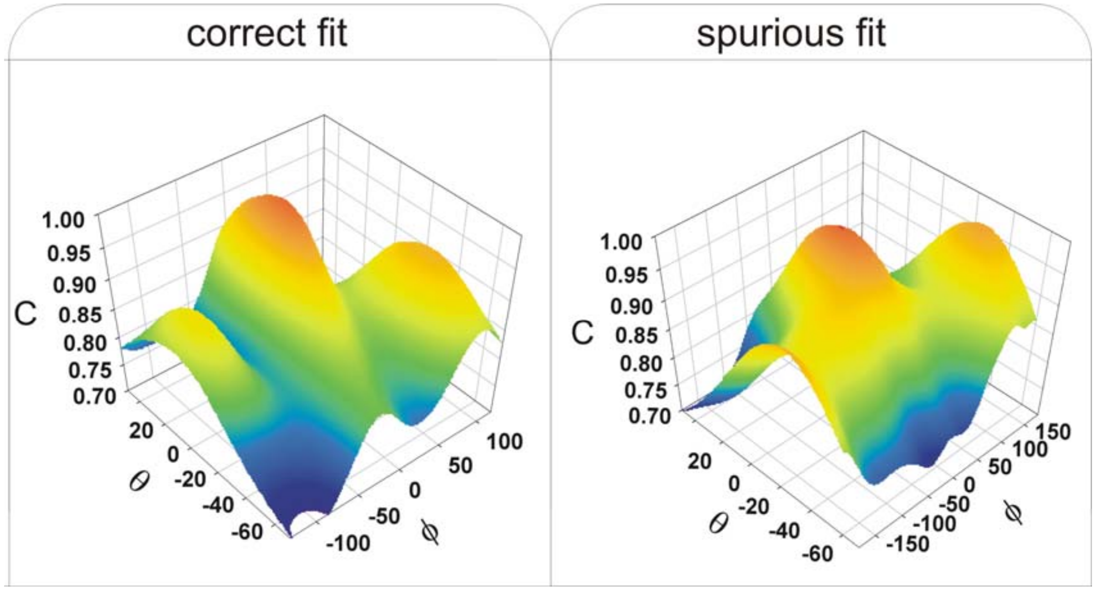
$$F^{-1}\left[F(\mathbf{e}\cdot\rho_{em})^* \cdot F(\mathbf{e}\cdot\rho_{calc})\right]$$

**UPDATE LATTICE:**
**SAVE MAX. *C* VALUE +**
**CORRESPONDING $\Theta,\Phi,\Psi$**

P. Chacón

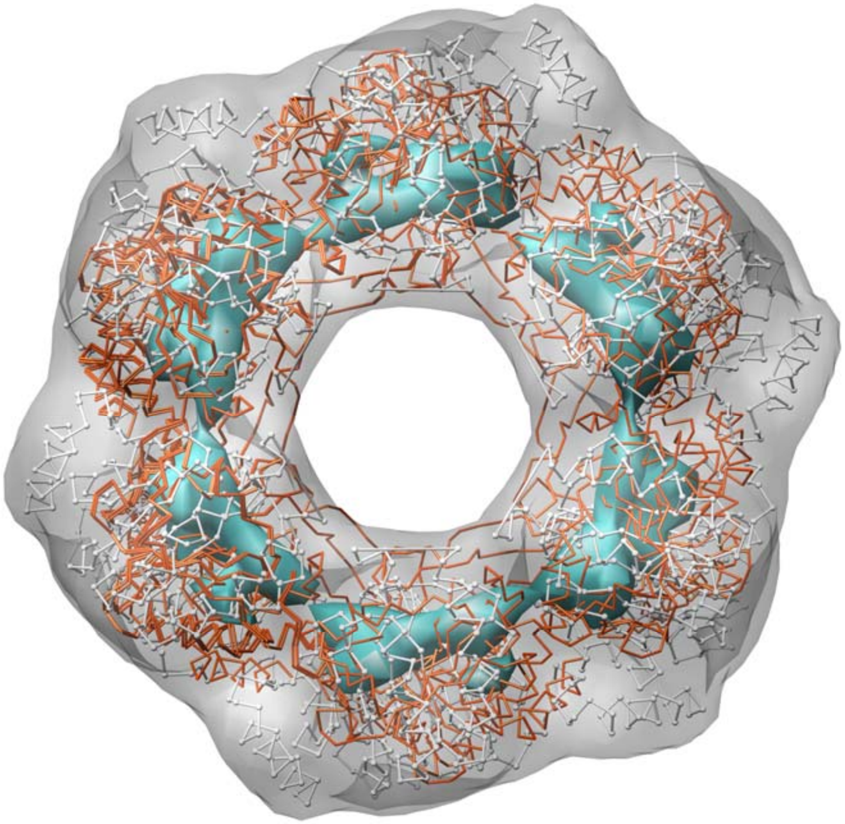
# Effect of Filter on Orientation

No Filter

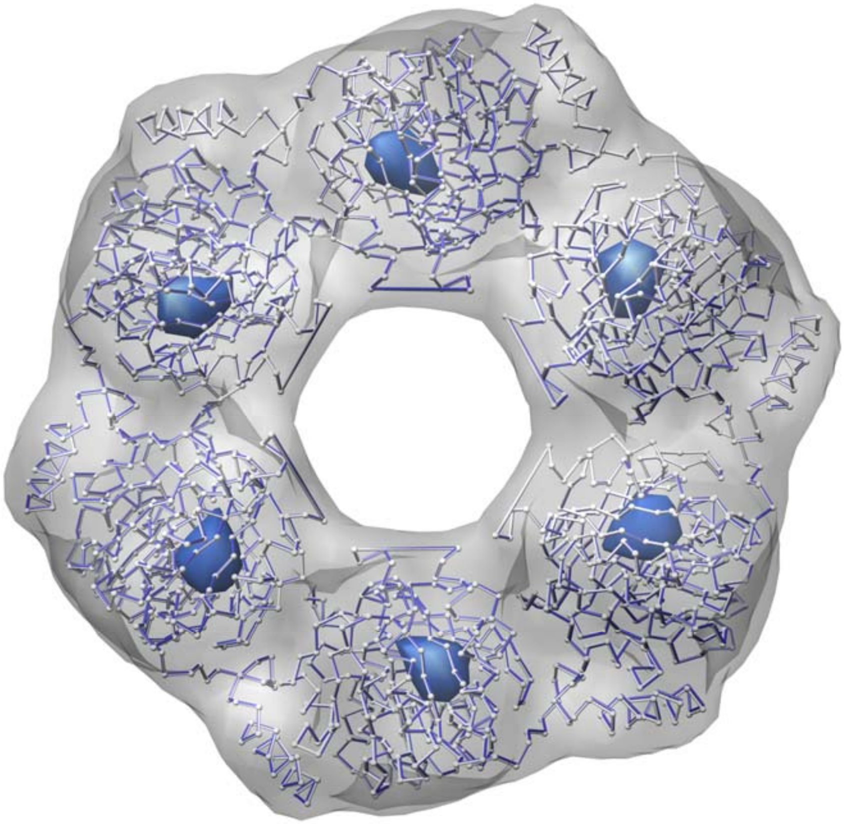


P. Chacón

# Example: RecA Translation Function



Standard cross-correlation

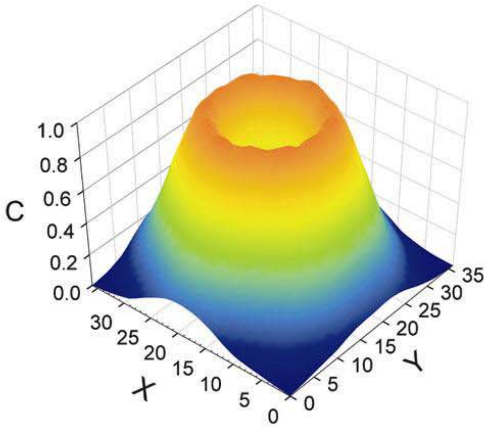


with Laplacian filtering

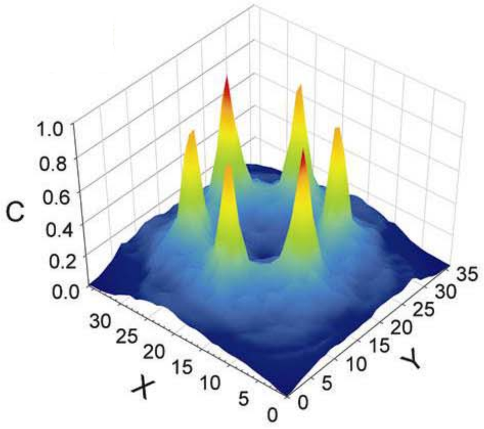
# Example: RecA

**Grid size 6Å**
**Resolution 15Å**
**9° steps (30481 rotations)**

standard cross-correlation



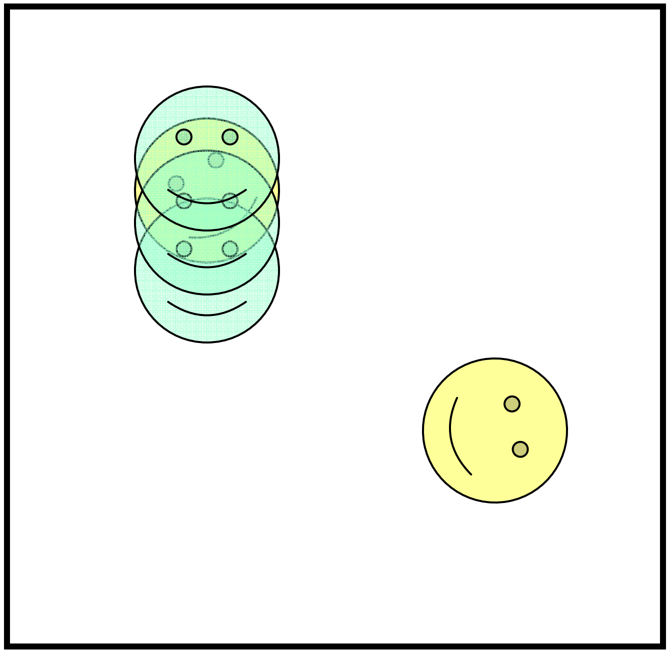
w/ Laplacian filtering



**Only Laplacian filtering successfully restores the initial position**

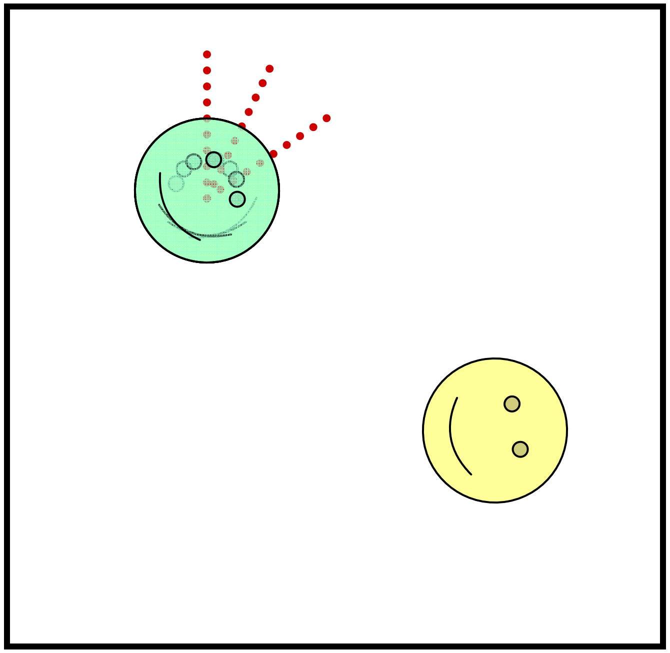
# Search Granularity

## Translational Granularity



Originates from voxel spacing

## Rotational Granularity



Originates from angular sampling

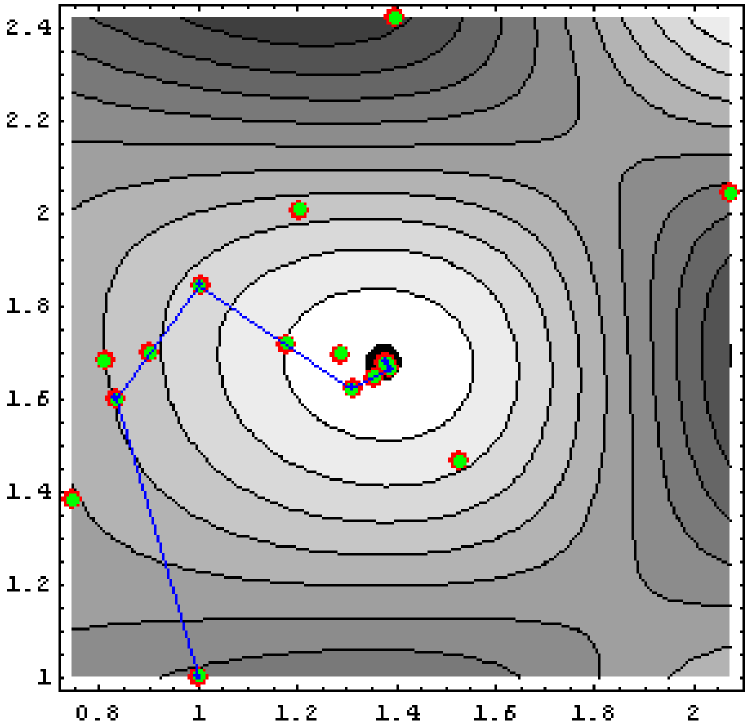
# Off-Lattice Refinement

The exhaustive search is limited to a grid of points in the 6D search space

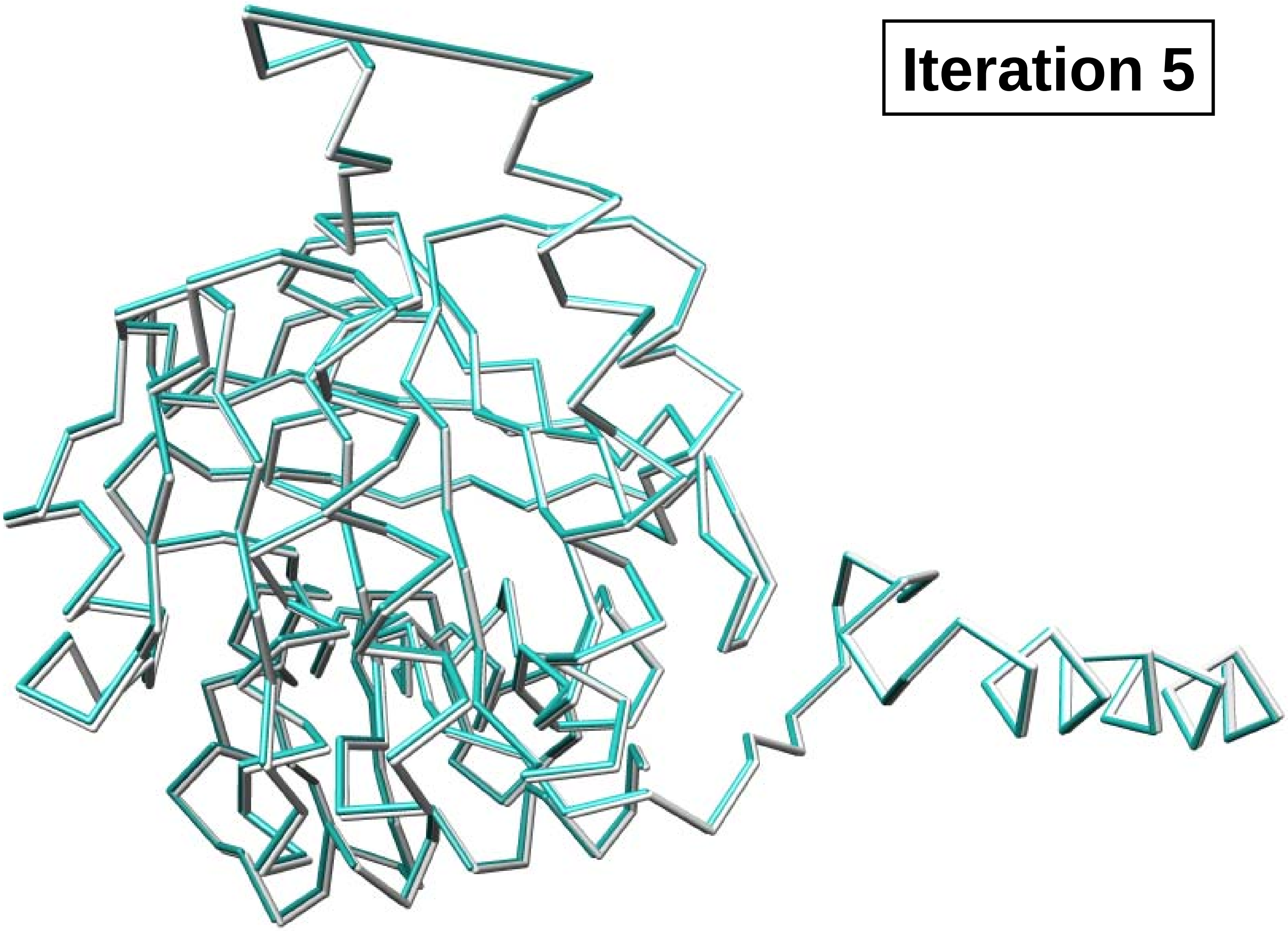
Improve the accuracy

Off-lattice (6D) local maximization of the correlation coefficient

Powell’s quadratically convergent maximization method can be used to perform a 6D search around the best fits found on the grid.

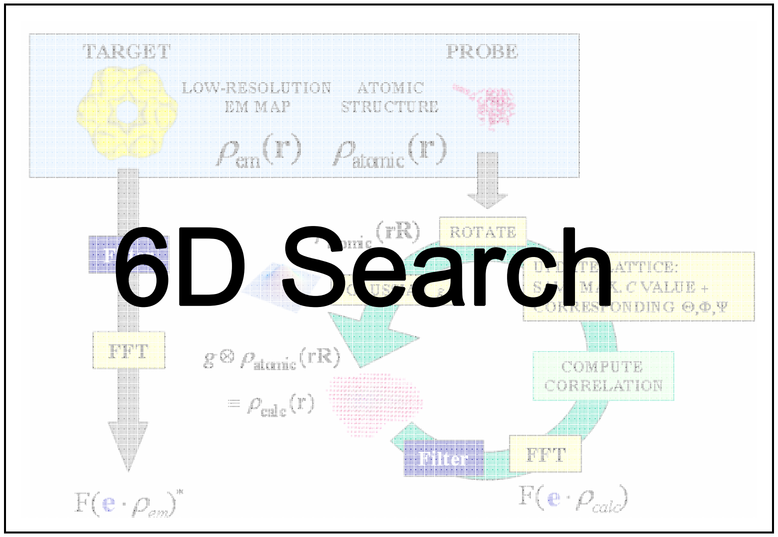


# Off-Lattice Refinement: Example

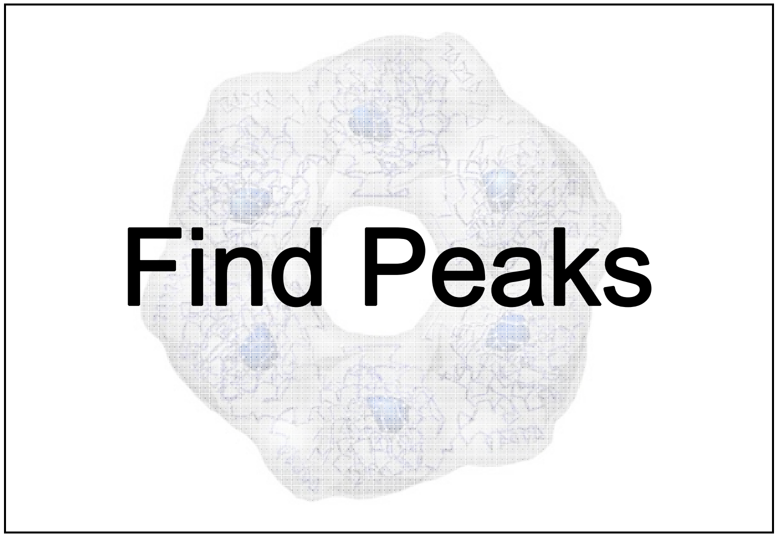


**Iteration 5**

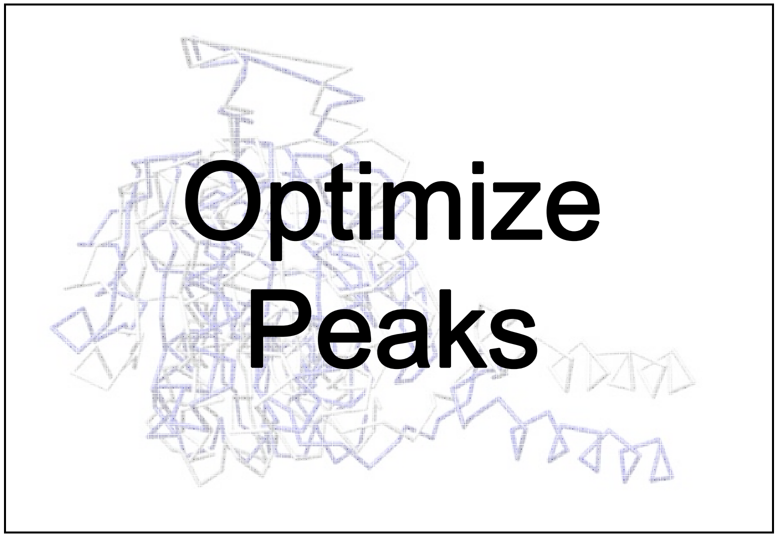
# Complete Workflow



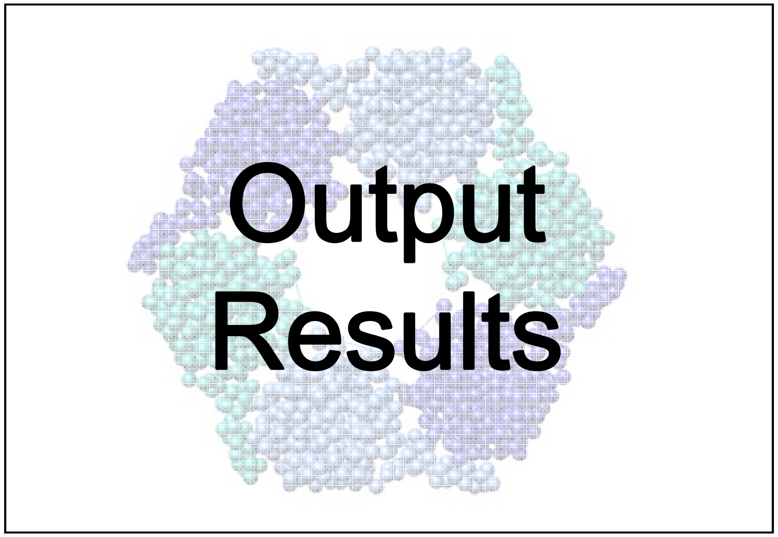
6D Search



Find Peaks

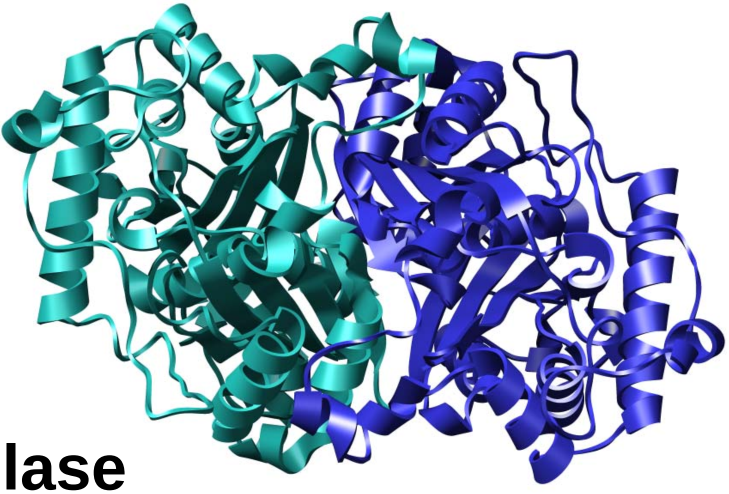


Optimize Peaks



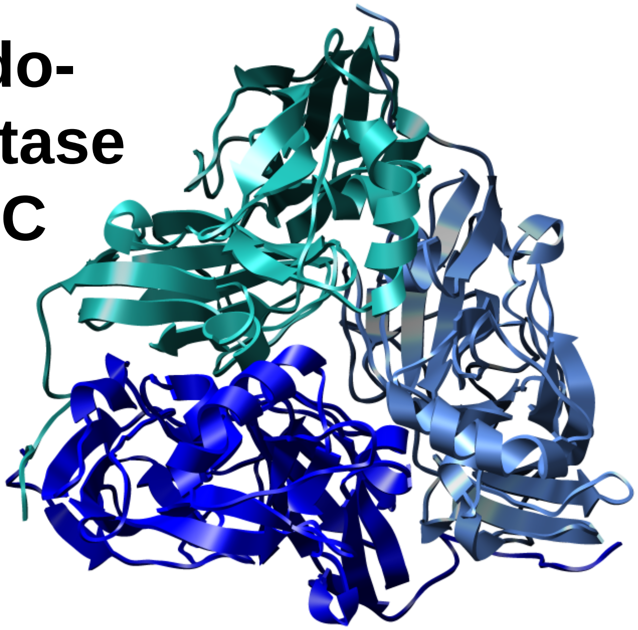
Output Results

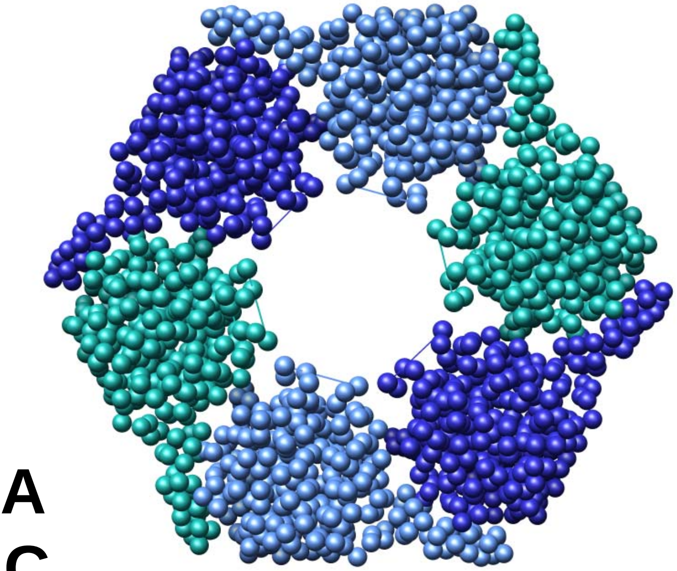
# Restoring Various Oligomers



**Thiolase**
**1AFW**

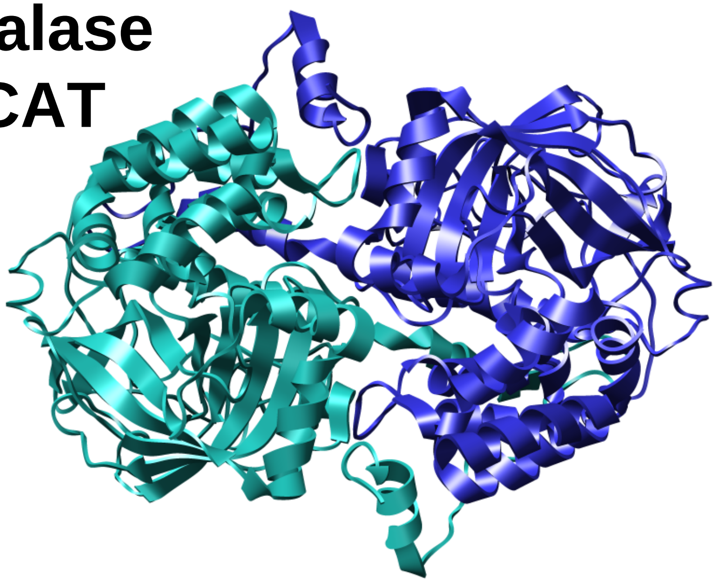
**Oxido-**
**reductase**
**1NIC**



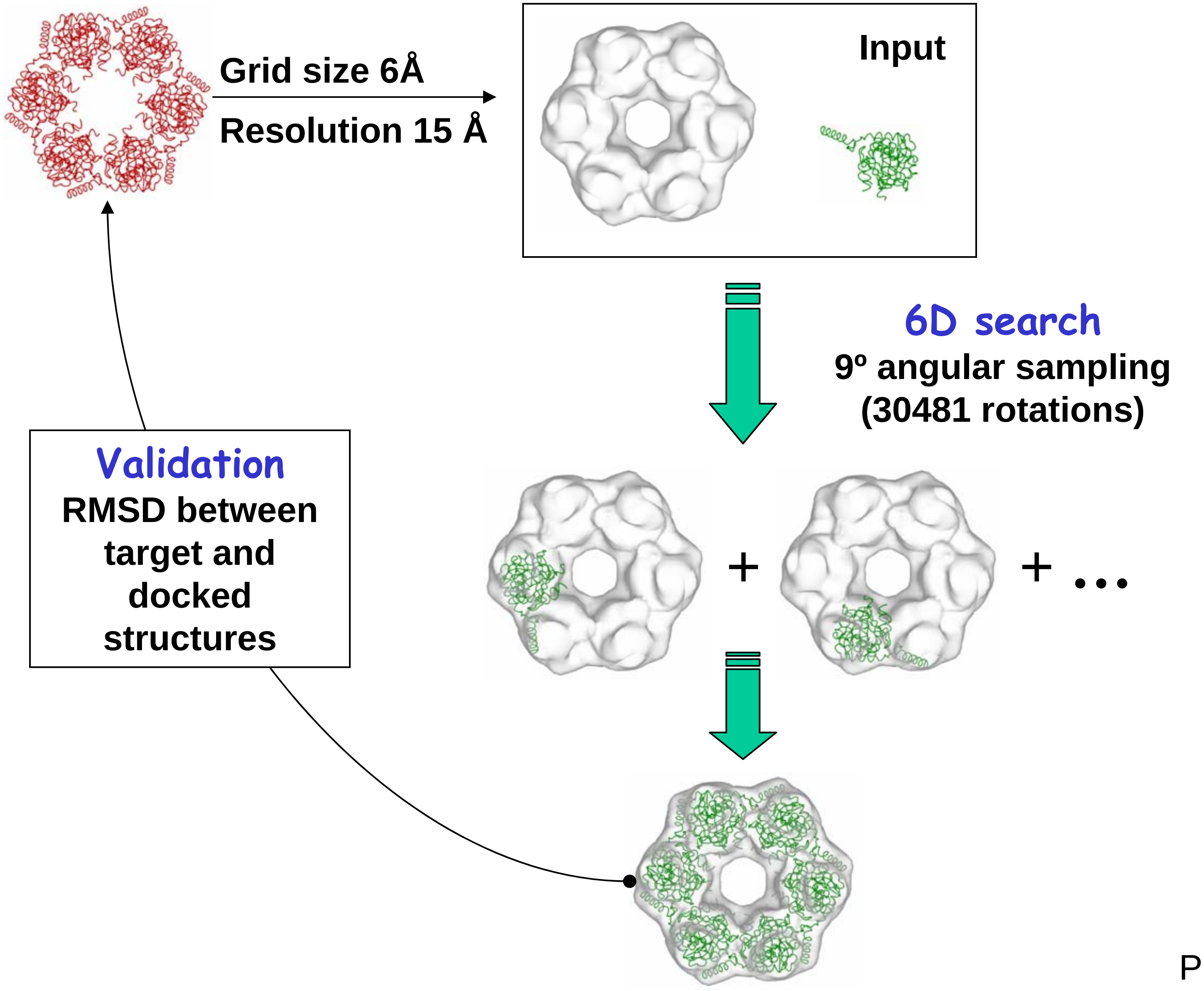


**RecA**
**2REC**

**Catalase**
**7CAT**



# Restoration Tests with Simulated Data



Grid size 6Å

Resolution 15 Å

**Input**

**6D search**

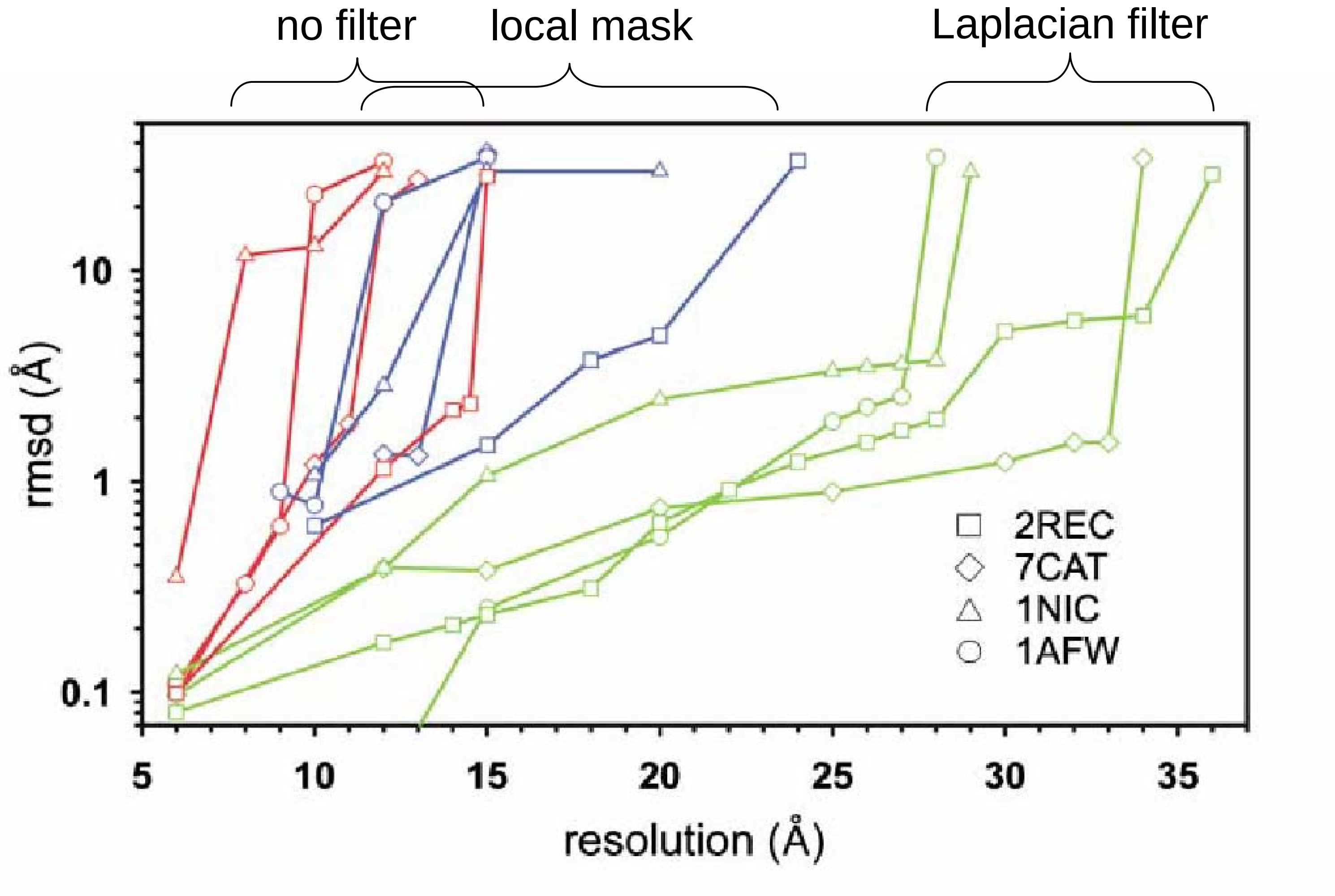
9º angular sampling
(30481 rotations)

**Validation**

RMSD between target and docked structures

P. Chacón

# Restoring Various Oligomers



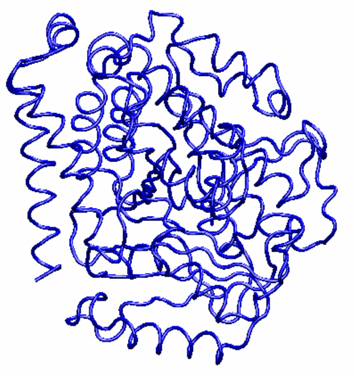
RecA (2REC), thiolase (1AFW), catalase (7CAT), and oxidoreductase (1NIC).

P. Chacón

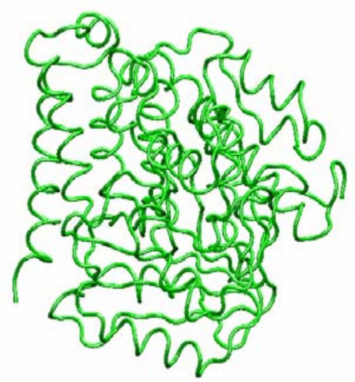
# Application to Microtubule Data

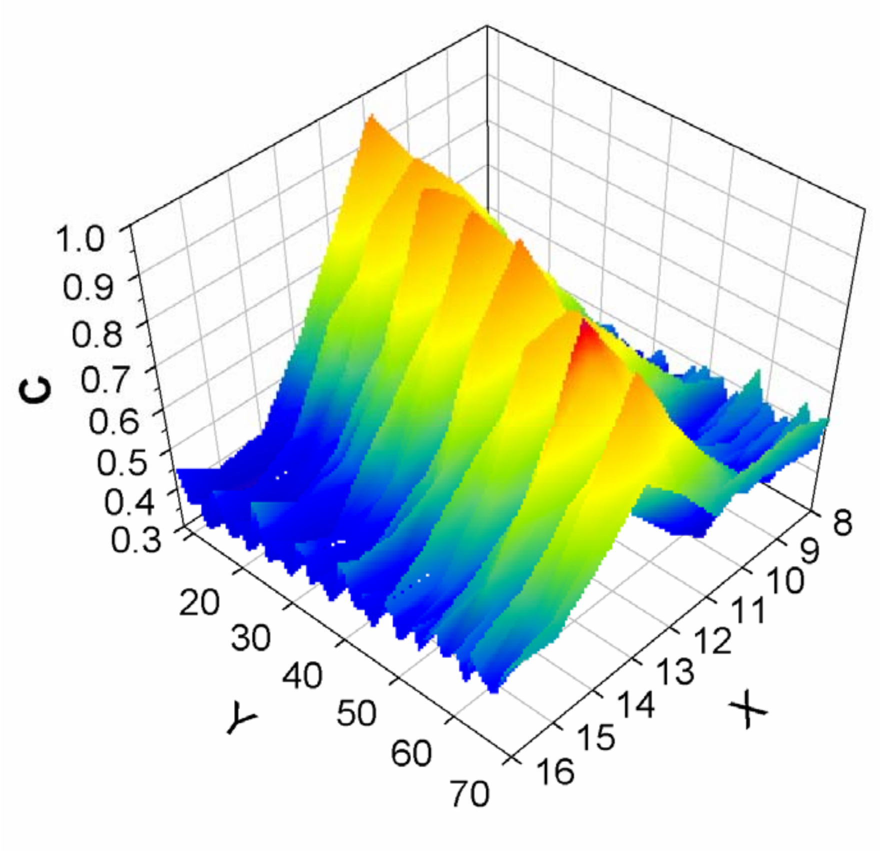
**Resolution 20Å**
**Angular sampling 9º**
**Grid size 5Å**

β-tubulin



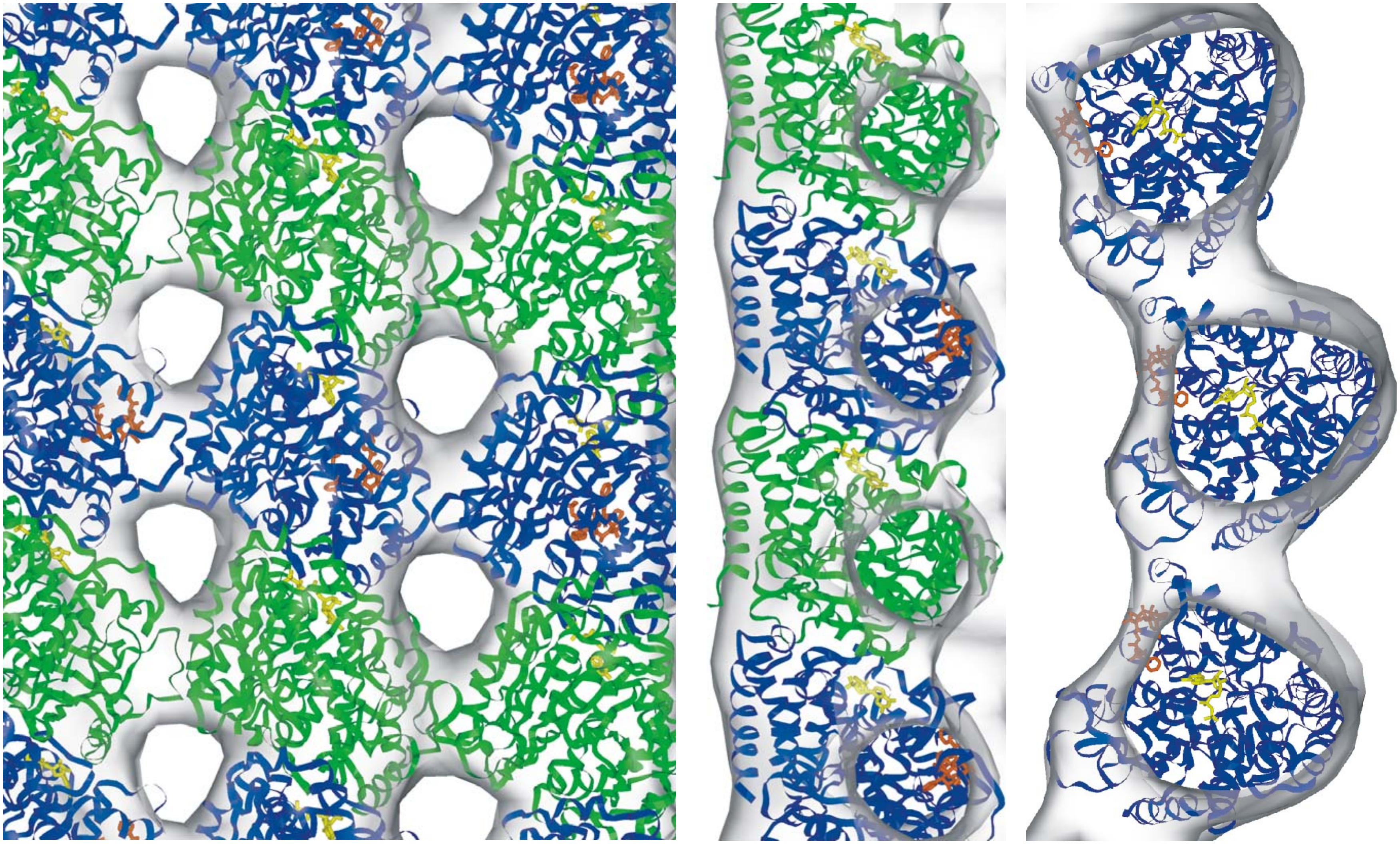
α-tubulin





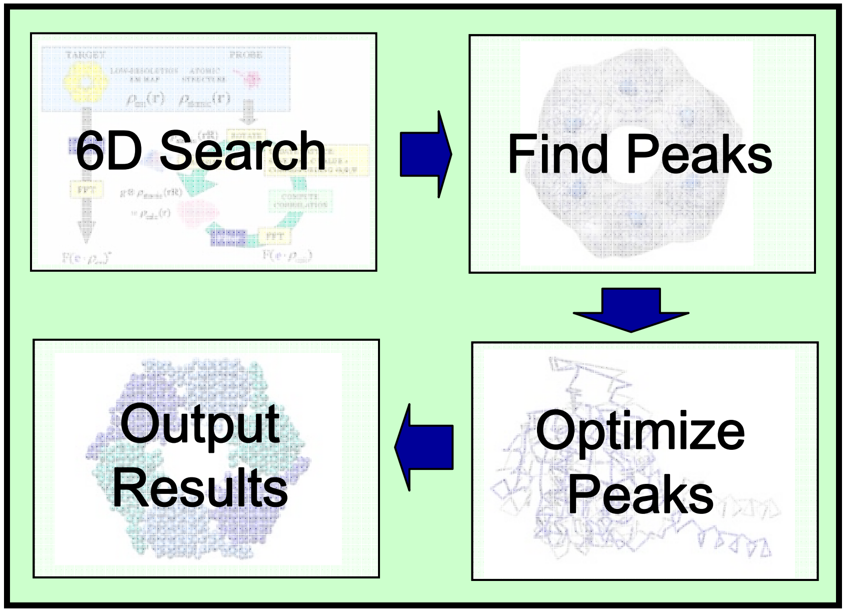
P. Chacón

# Microtubule Model



P. Chacón

# The Situs Software Package

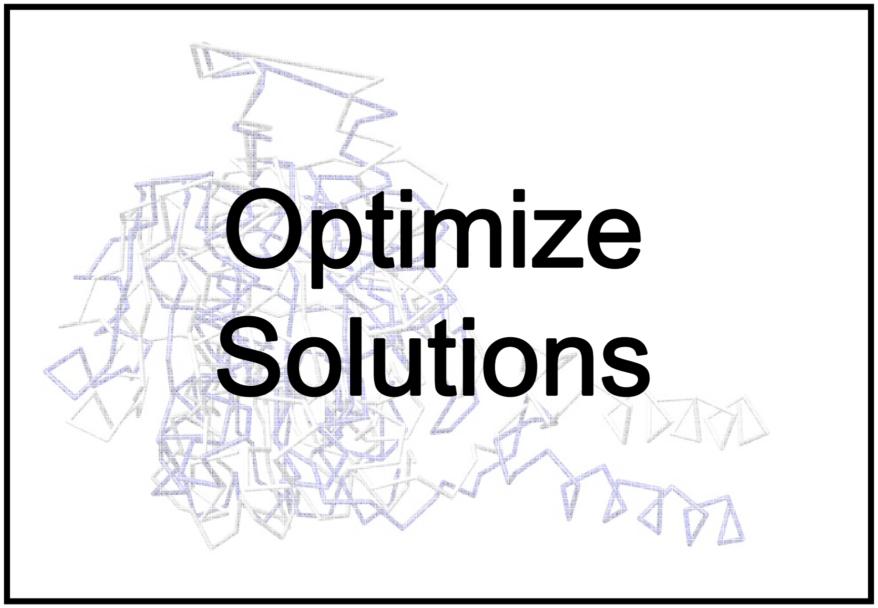


6D Search

Find Peaks

Optimize Peaks

Output Results



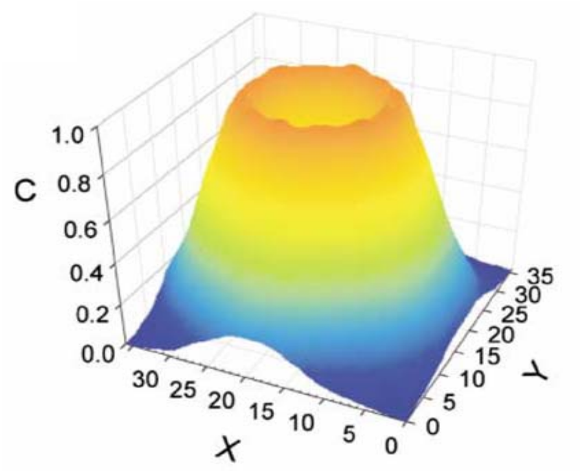
Optimize Solutions

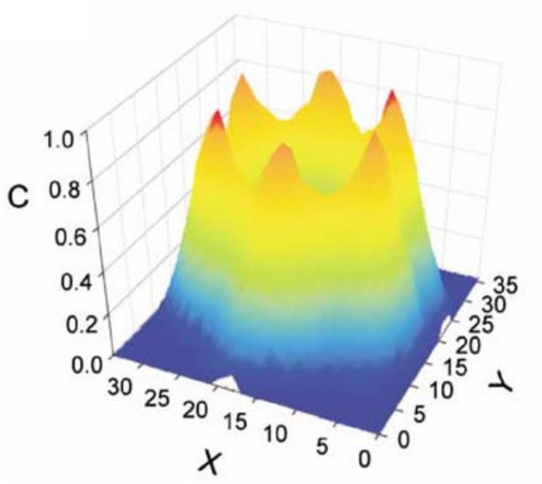
Colores

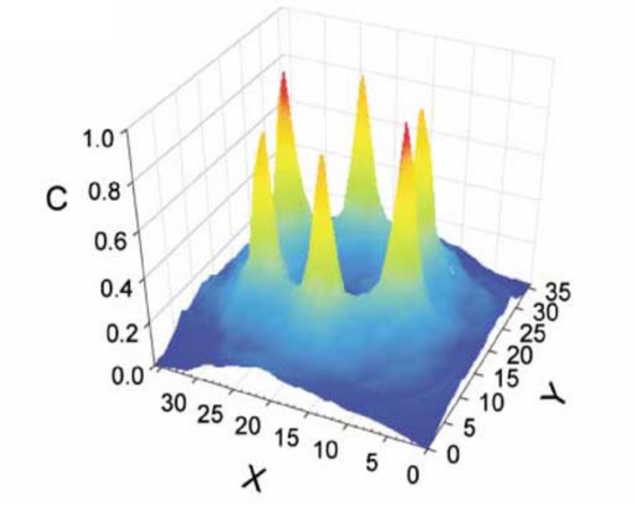
Colacor

**Tutorial: Thursday afternoon**

# Summary: Correlation Based Matching

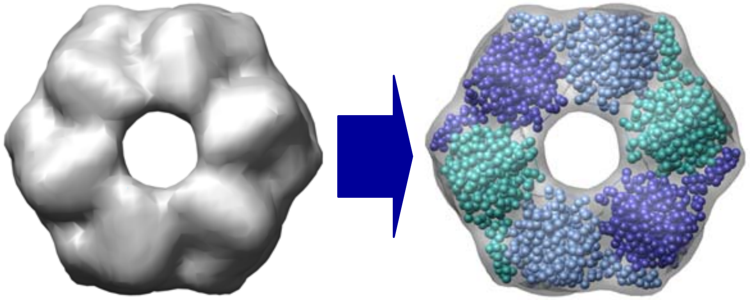

**No filter**


**Local mask**


**Laplacian filter**

→ *Increasing Fitting Contrast* →

**Situs** 6D exhaustive searches:



- Rigid Body
- Density Filtering
- Fast Translational Matching
- Fast Rotational Matching (later talk)

# Resources and Acknowledgements

WWW:

http://situs.biomachina.org

http://situs.biomachina.org/tutorial_colores.html

Papers:

http://situs.biomachina.org/fref.html

Acknowledgement:

Pablo Chacón, Valerio Mariani, Paul Boyle,
Willy Wriggers