Visualization of pinfold simulations

Sebastian Pötzsch

Faculty of Mathematics and Computer Science
University of Leipzig

Overview

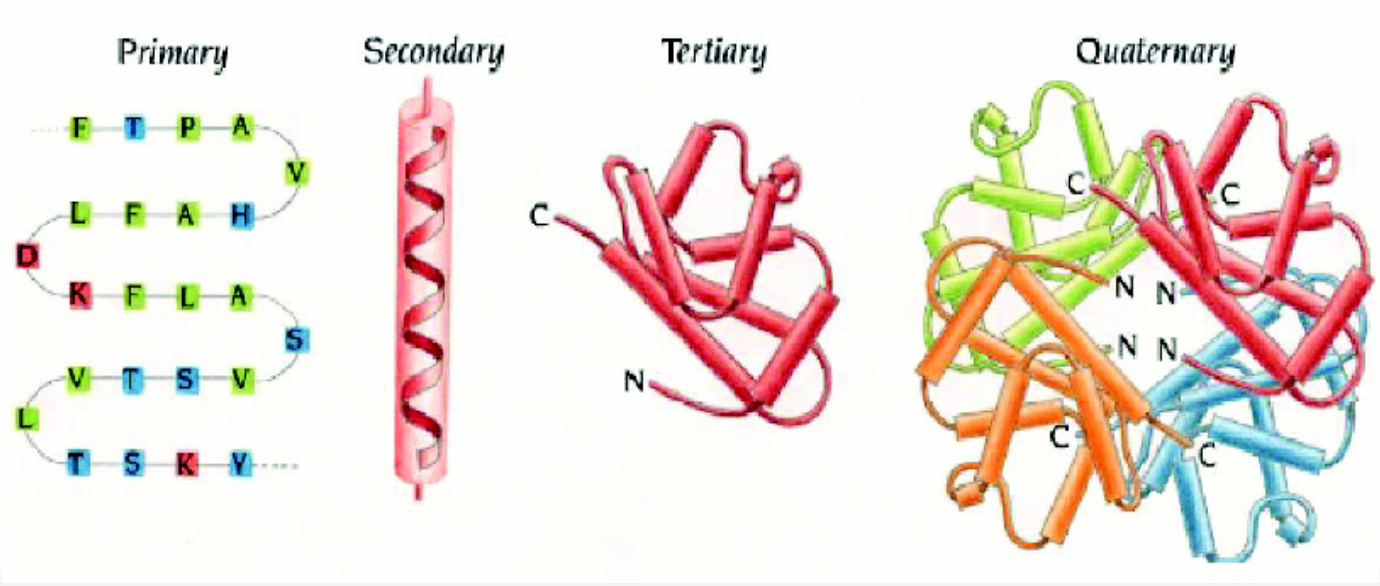
1. Introduction
   - 1.1 Protein folding problem
   - 1.2 HP-model
   - 1.3 Pinfold simulation
2. Information Visualization
   - 2.1 Shneiderman’s mantra
   - 2.2 Overview
   - 2.3 Zoom and filter
   - 2.4 Detail views
3. Pinfoldvis Demo

1. Introduction

- Proteins:
  - Complex macromolecules
  - Chains of amino acids linked by peptide bonds
  - Typical a few hundred amino acids
  - Essential functions in organisms
  - Function depends on the structure
  - Complex structures

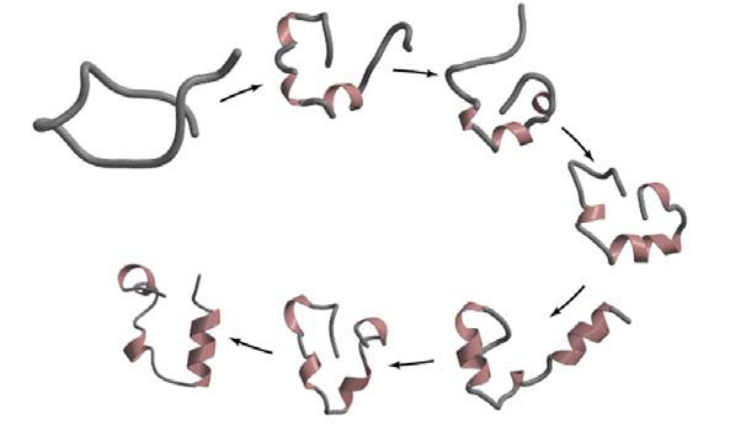
1. Introduction

- Protein structure:
  - Primary: amino acid sequence
  - Secondary: local conformations (α helices, β sheets, β turns)
  - Tertiary: real three dimensional structure
  - Quaternary: connection of polypeptide chains



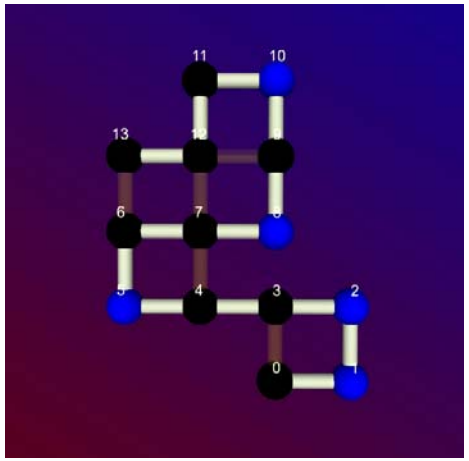
1.1 Protein folding problem

- Protein folding:
  - Problem: Given amino acid sequence → native structure
  - Not completely understood
  - Important to understand
  - Known forces (hydrogen bonds, hydrophobic interactions, Van der Waals force, disulfide bonds)



1.2 HP-Model

- By Dill and Lau, 1989
  - Amino acids → beads, unit size
  - Bonds → straight sticks, unit length
  - Positions fixed to lattice points
  - Two classes H = {A,C,I,L,M,F,W,Y,V} and P = {R,N,D,E,Q,G,H,K,P,S,T}
  - Simple energy function



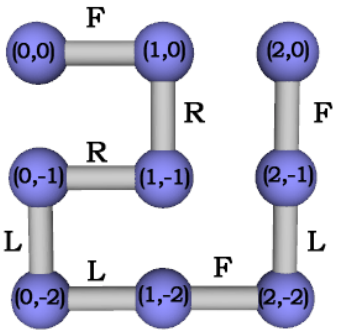
- HPNX-Model
  - H (hydrophobic), P (positive), N (negative), X (uncharged)

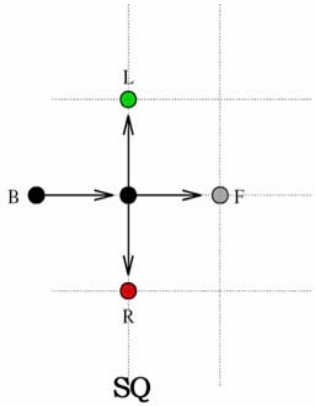
	H	P
H	-1	0
P	0	0

	H	P	N	X
H	-4	0	0	0
P	0	1	-1	0
N	0	-1	1	0
X	0	0	0	0

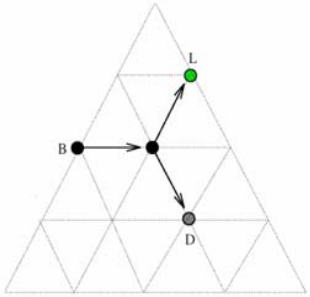
1.3 Pinfold simulation

- Based on HP/HPNX-Model
- Simulates folding on lattices
- Structure description with self-avoiding walks (SAW)

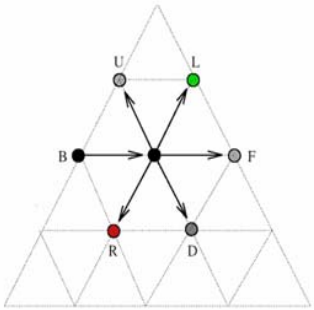




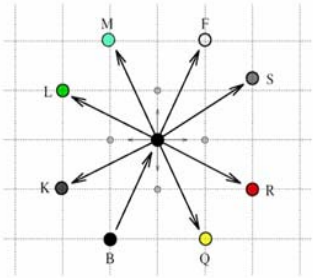
SQ



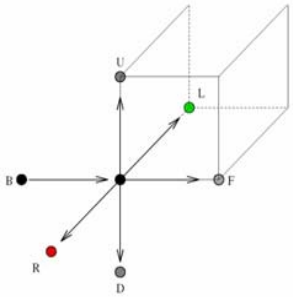
HEX



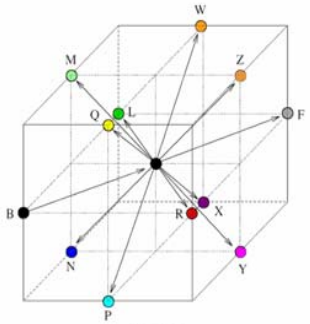
TRI



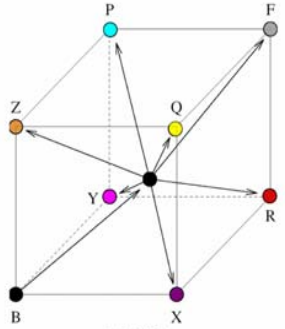
KM



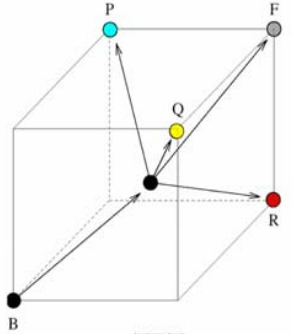
SC



FCC



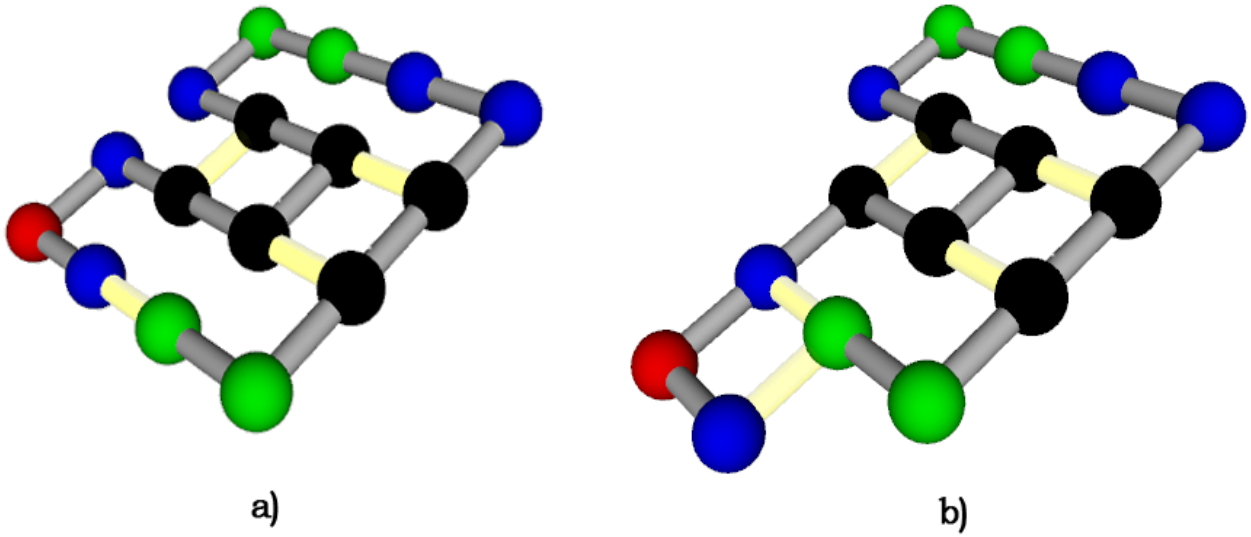
BCC



TET

1.3 Pinfold simulation

- Example
  - Model: HPNX
  - Lattice: SQ
  - Sequence: NNHHPPNNPHHHHPXP



a) Structure: FLFFLFFLLFRRFLL
Energy: -13

b) Structure: FLFFLFFLLFRRLFL
Energy: -14

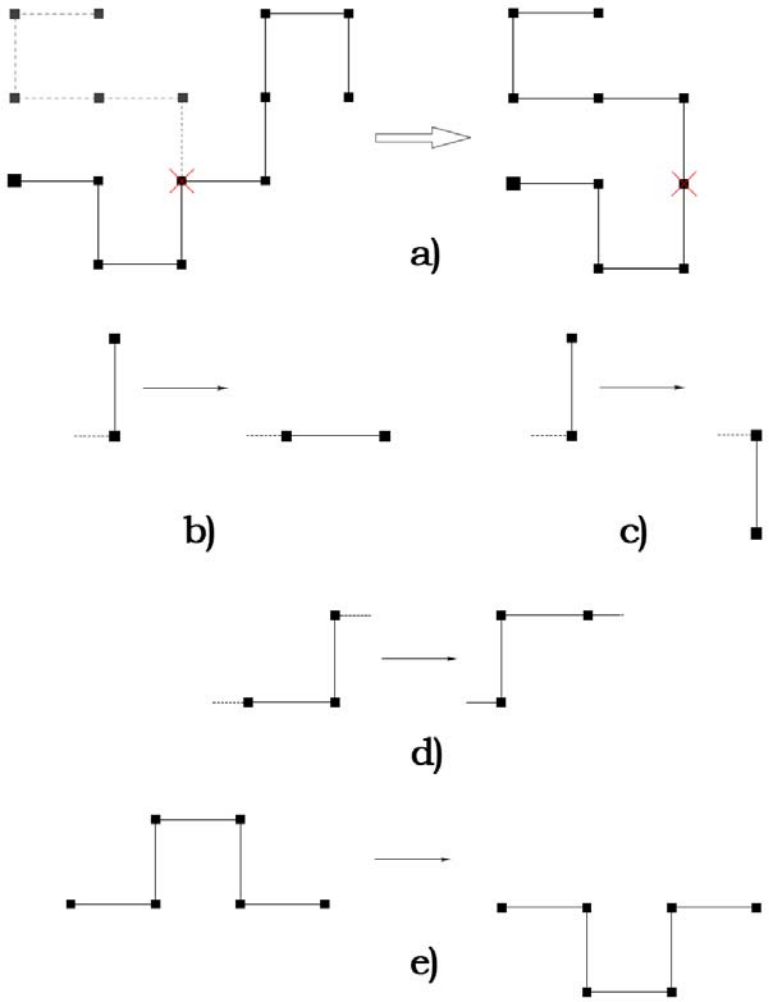
1.3 Pinfold simulation

- Move set
  - Set of rules
  - Moves have to be reversible
  - Results have to be in conformation space
  - Ergodicity
- describes topology of conformation space
- a) pivot move

  b) c) end move

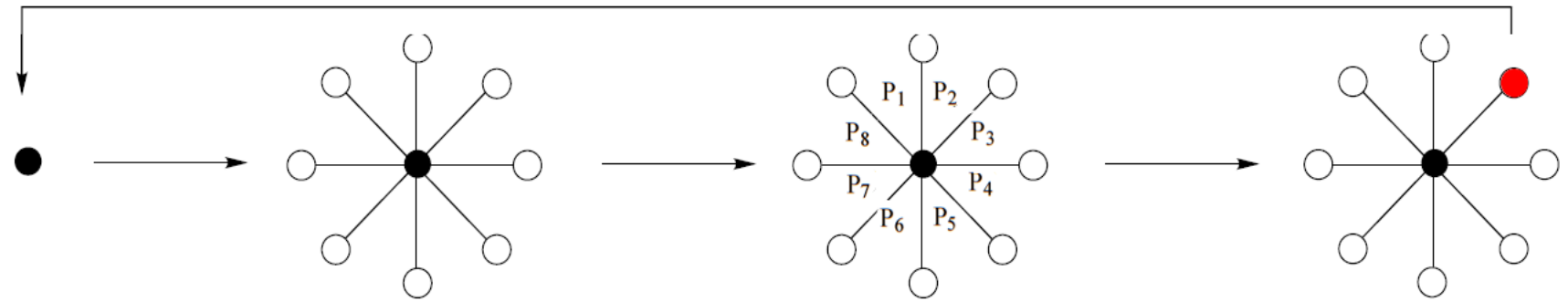
  d) corner move

  e) crankshaft-move



1.3 Pinfold simulation

- Algorithm:
  1. Begin with start structure.
  2. Check terminating conditions
     (max. time reached, stop structure found).
  3. Generate all valid neighbor structures with their energy.
  4. Compute transition rates. $p_i = \min\{1, \exp(-\frac{\Delta E}{kT})\}$
  5. Choose neighbor structure with probability proportional to p.
  6. Advance clock $\frac{1}{\sum_i p_i}$ → Step 2



1.3 Pinfold simulation

- Output:
  - Sequence
  - Start structure
  - Chosen structure, Energy, Elapsed time
- **Our task:**
  - Visualize pinfold output
  - Support data analysis
  - Emphasize possible relationships
  - Provide simulation comparison

```
HHPHPHPPHPHPHPH
FFLRLRFLLFFLRL
FFFRLRFLLFFLRL  -4.00      0.056
FFFRLRFLLFFLRF  -3.00      0.735
FFFFRRFLLFFLRF  -2.00      2.520
FRFFLRFLLFFLRF  -4.00      2.849
FRFFLRFLLFFLRF  -4.00      3.521
FRFRLRFLLFFLRF  -3.00      3.689
FLFRLRFLLFFLRL  -4.00      4.444
FRFRLRFLLFFLRR  -3.00      4.587
FRFRLRFRLFFLRR   0.00      5.556
FRFRLRFLLFFLRR  -3.00      5.908
FLFRLRFLLFFLRL  -4.00      8.026
FFFRLRFLLFFLRL  -4.00      8.624
FFLRLRFLLFFLRF  -3.00      9.211
FFRRLRFLLFFLRL  -4.00      9.381
FFLRLRFLLFRLRL  -1.00      9.786
FFLRLRFLLRFLRL   0.00     11.068
FFLRFRFLLFFLRL  -2.00     12.552
...
```

2. Information Visualization

- “to visualize”: form a mental vision, image, or picture of something not visible or present to sight, or an abstraction; to make visible to the mind or imagination. (The Oxford English Dictionary; 1989)

- Task:
  - Provide data analysis and exploration
  - Emphasize relationships
  - Uncover regularities
  - Expose the unseen (abstract)
  - Speedup cognition

2.1 Shneiderman’s mantra

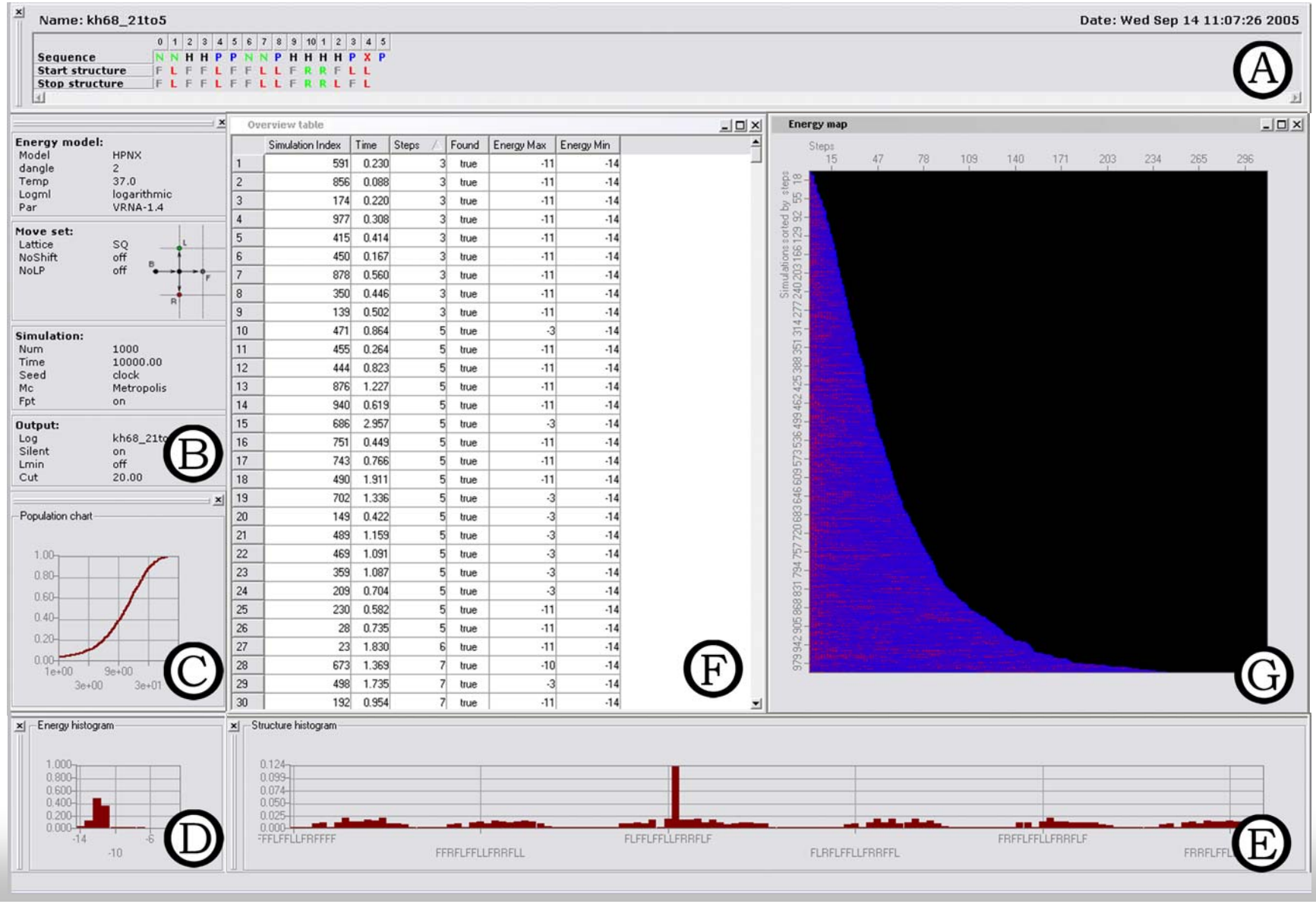
“Overview first, zoom and filter, details on demand.”

Idea:

- Start with an overview for a better orientation
- Let user decide to filter interesting data
- Show different details only on demand for selected data

2.2 Overview

"Overview first, zoom and filter, details on demand."

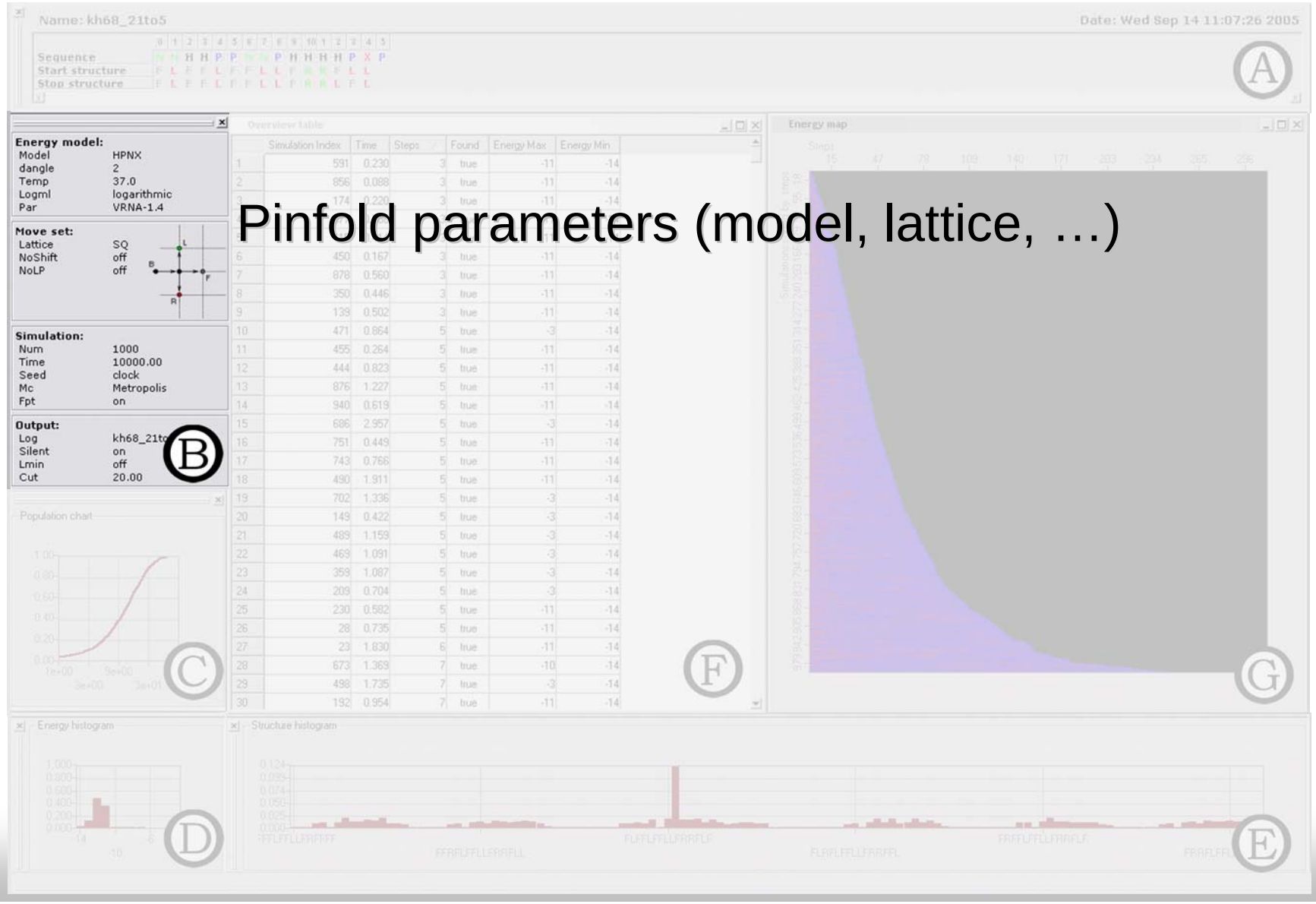


2.2 Overview

“Overview first, zoom and filter, details on demand.”

2.2 Overview

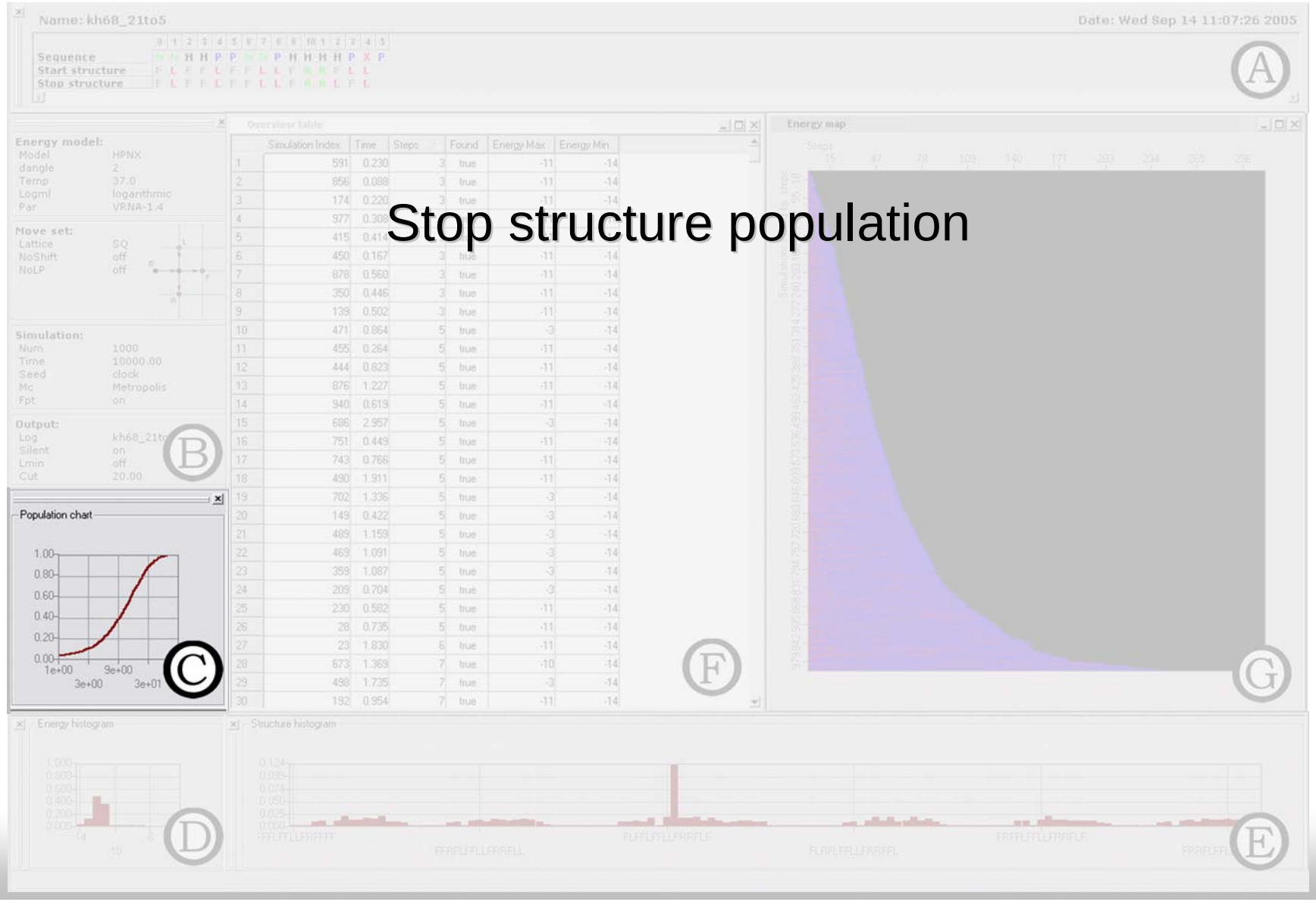
"Overview first, zoom and filter, details on demand."



Pinfold parameters (model, lattice, …)

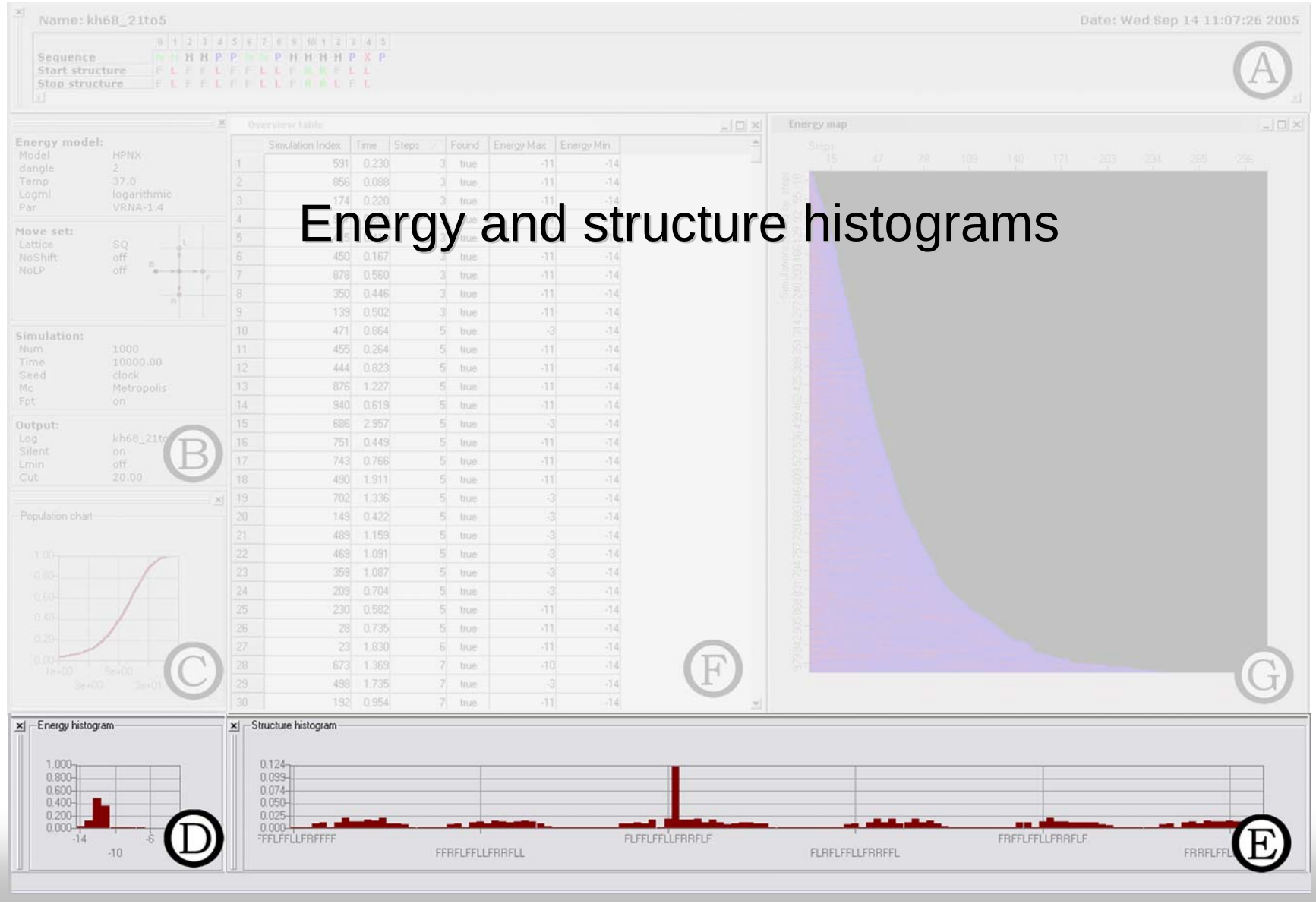
2.2 Overview

"Overview first, zoom and filter, details on demand."



2.2 Overview

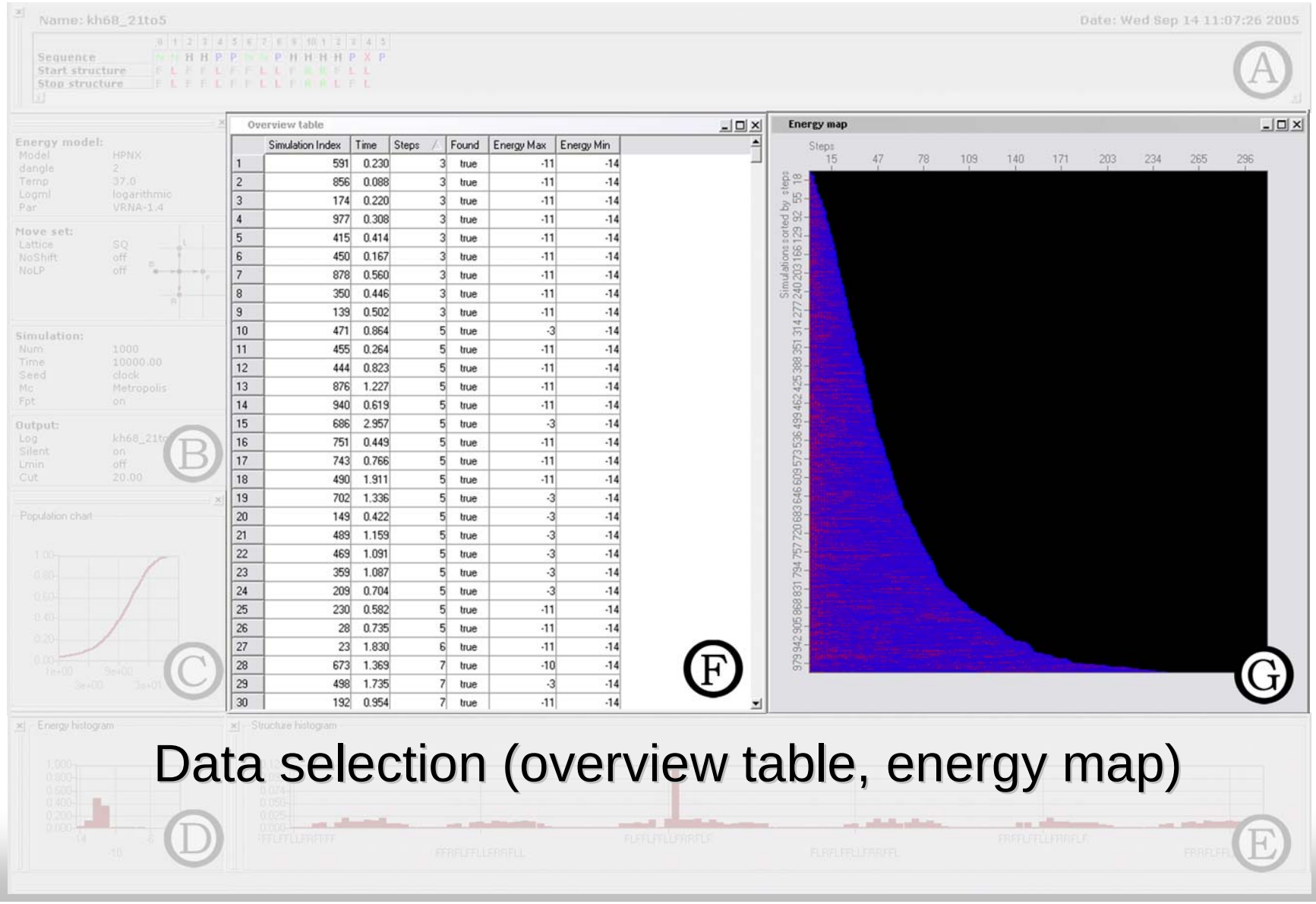
"Overview first, zoom and filter, details on demand."



Energy and structure histograms

2.3 Zoom and filter

"Overview first, zoom and filter, details on demand."

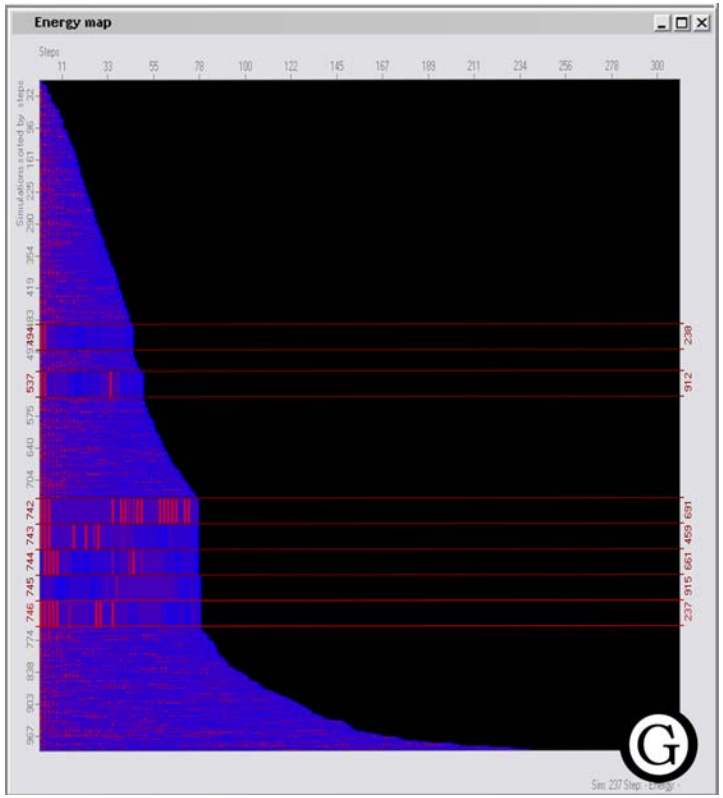


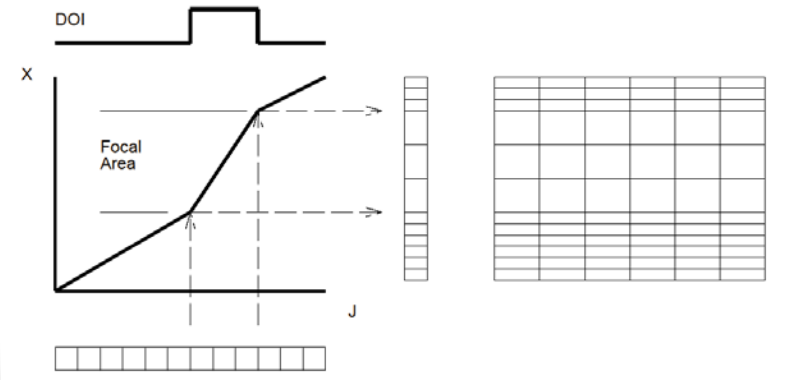
Data selection (overview table, energy map)

2.3 Zoom and filter

"Overview first, zoom and filter, details on demand."

- Energy map
- Focus+context technique
- Huge data sets ↔ limited screen size
- Details and overview in one view
- Hold up users attention as long as possible





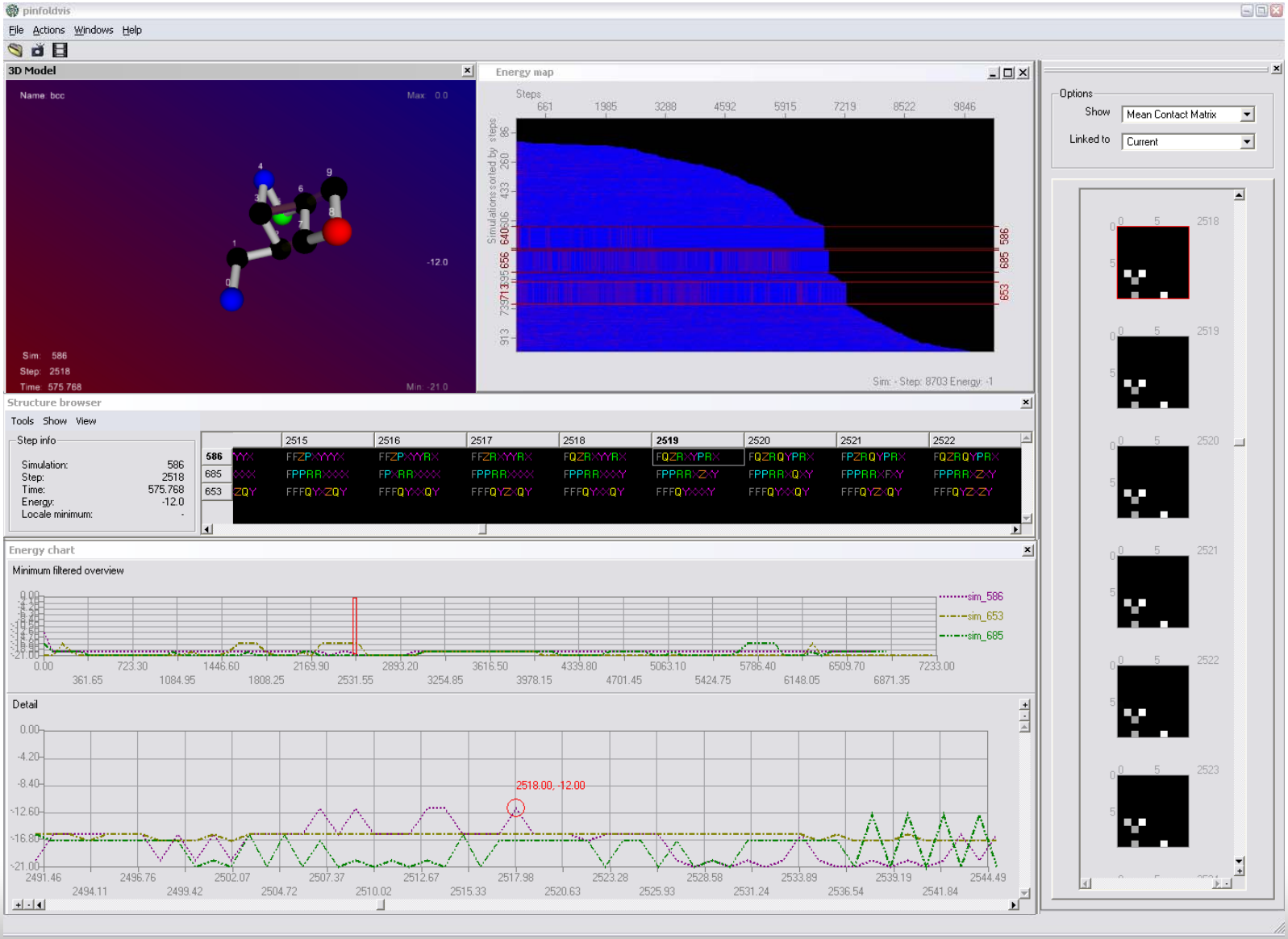
2.4 Detail views

"Overview first, zoom and filter, details on demand."

- Multiple Views
  - Two or more distinct views
  - Support the investigation of a single conceptual entity
  - Views can differ in data or visual representation
  - Different perspectives
  - Allows direct comparison in real-time
  - Linking several views emphasizes relationships

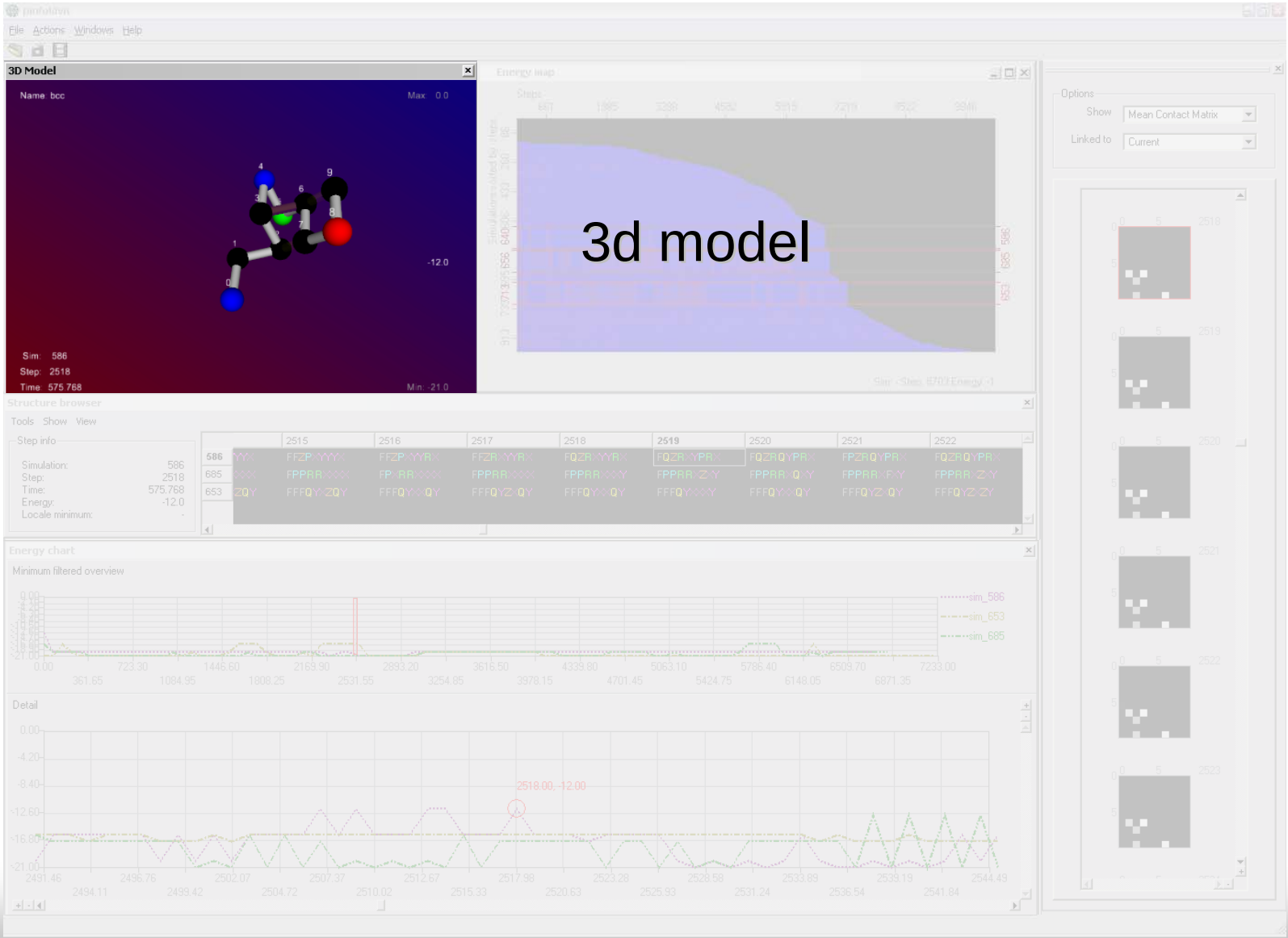
2.4 Detail views

“Overview first, zoom and filter, details on demand.”



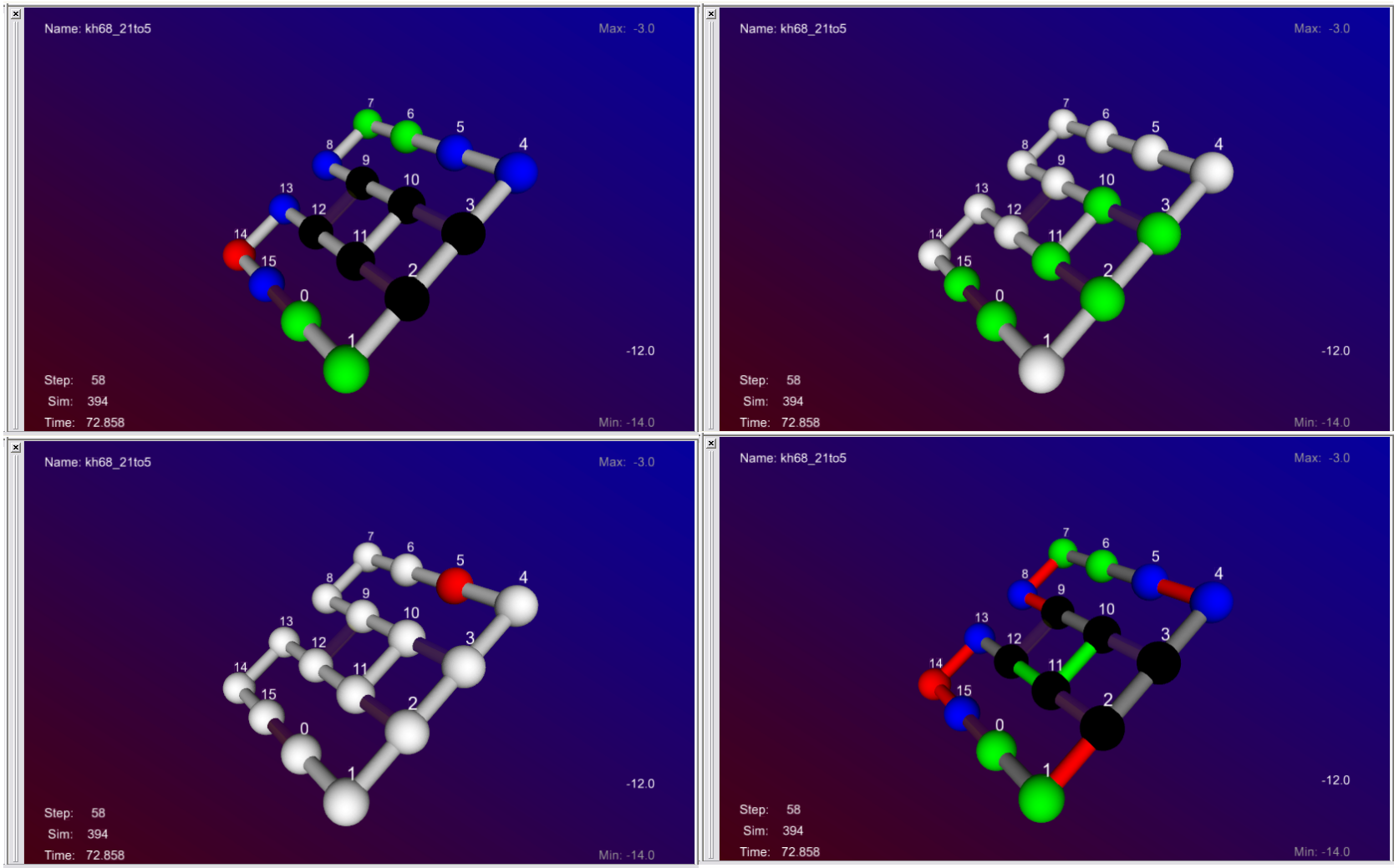
2.4 Detail views

“Overview first, zoom and filter, details on demand.”



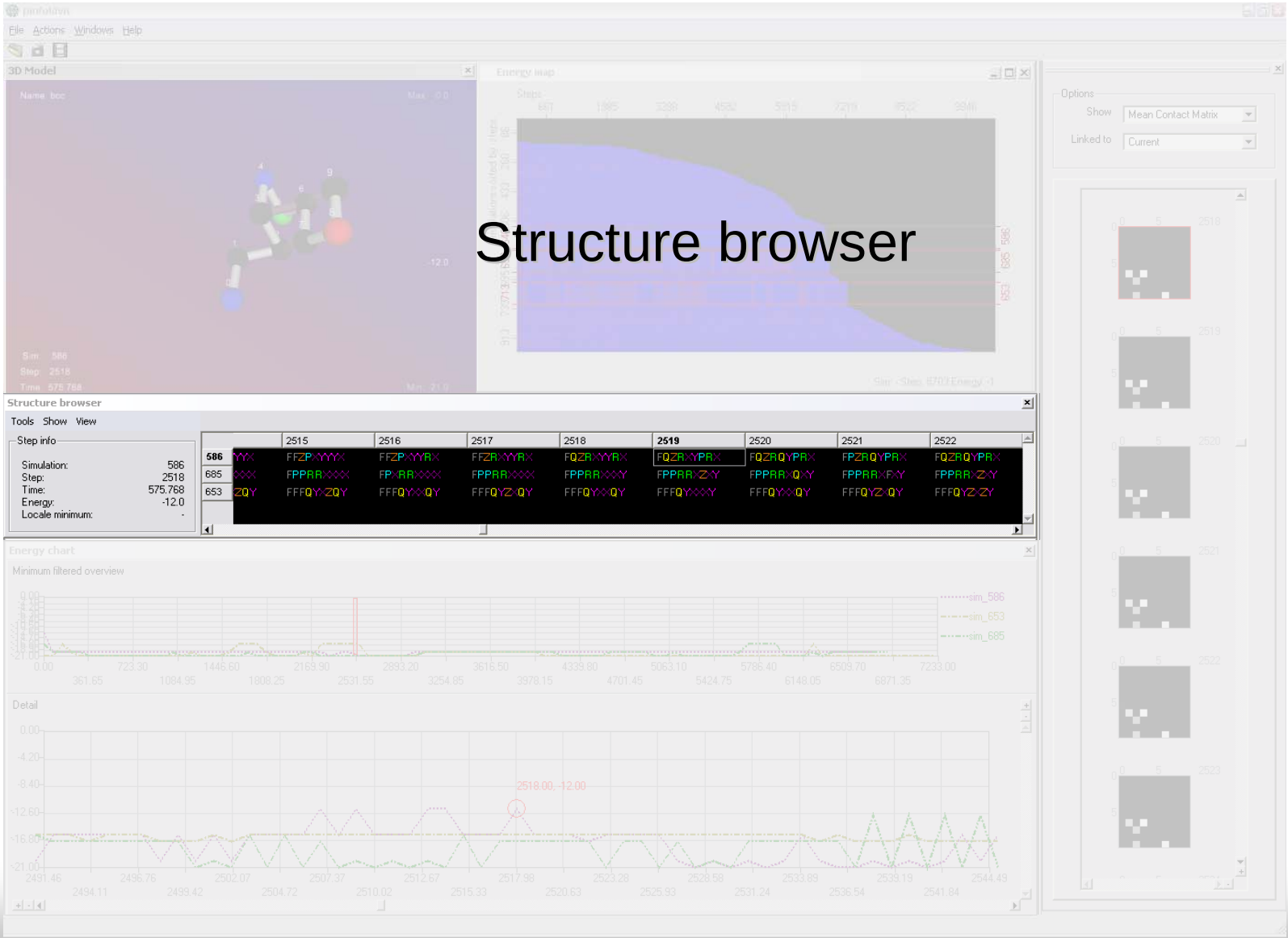
2.4 Detail views

- 3D structure viewer



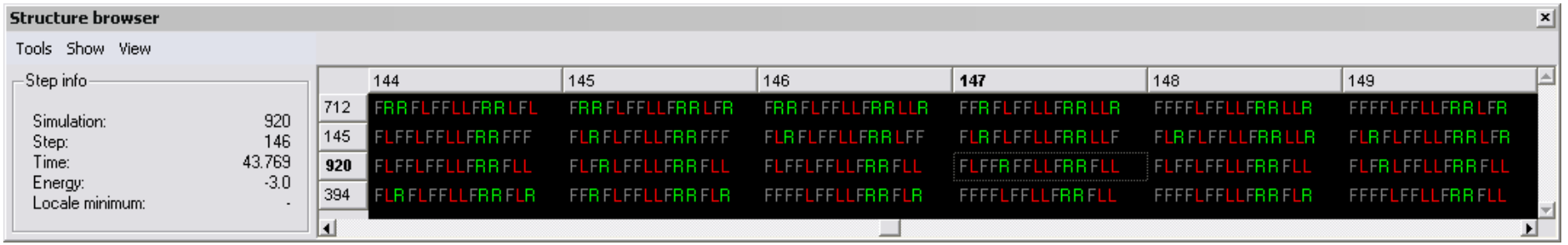
2.4 Detail views

“Overview first, zoom and filter, details on demand.”



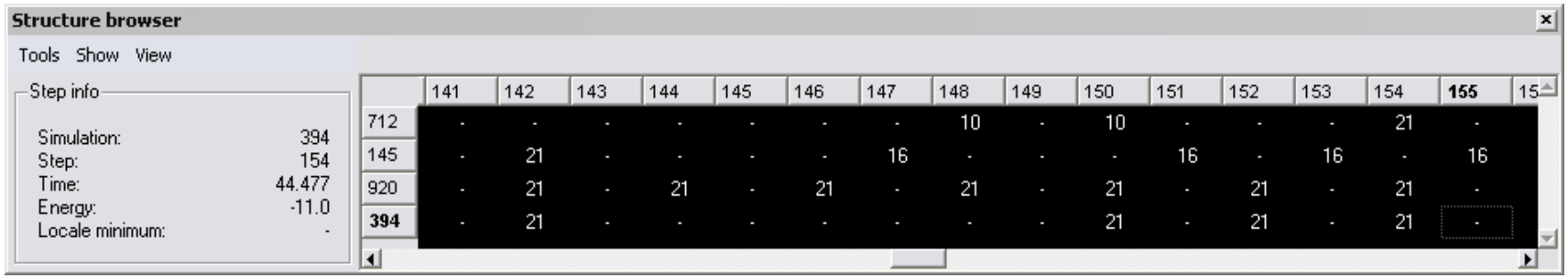
2.4 Detail views

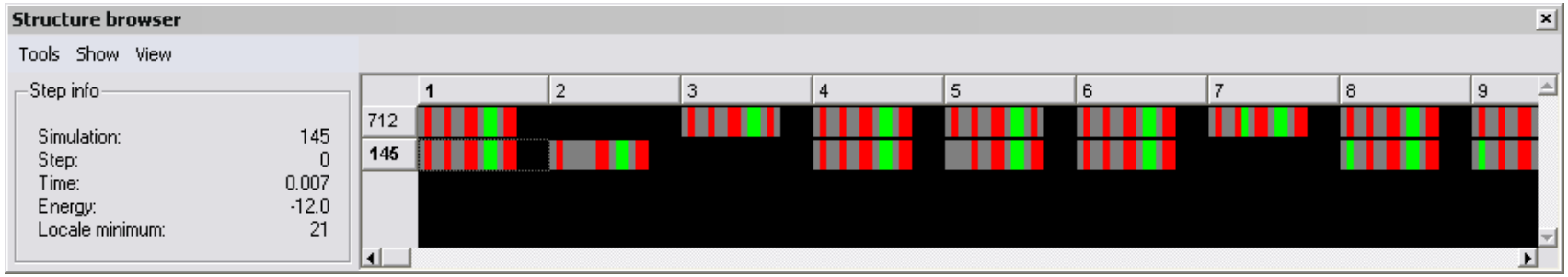
- Structure browser



2.4 Detail views

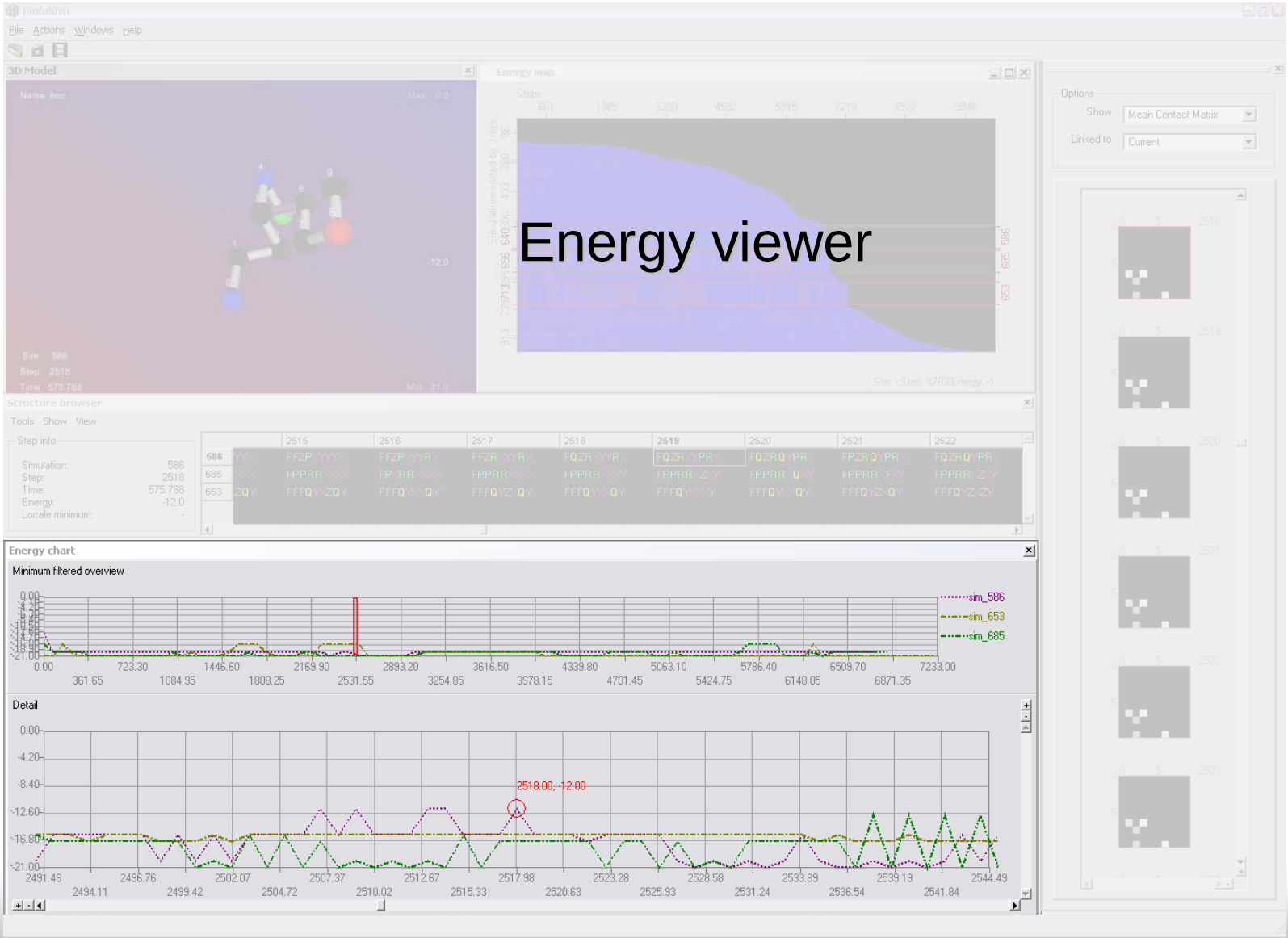
- Structure browser





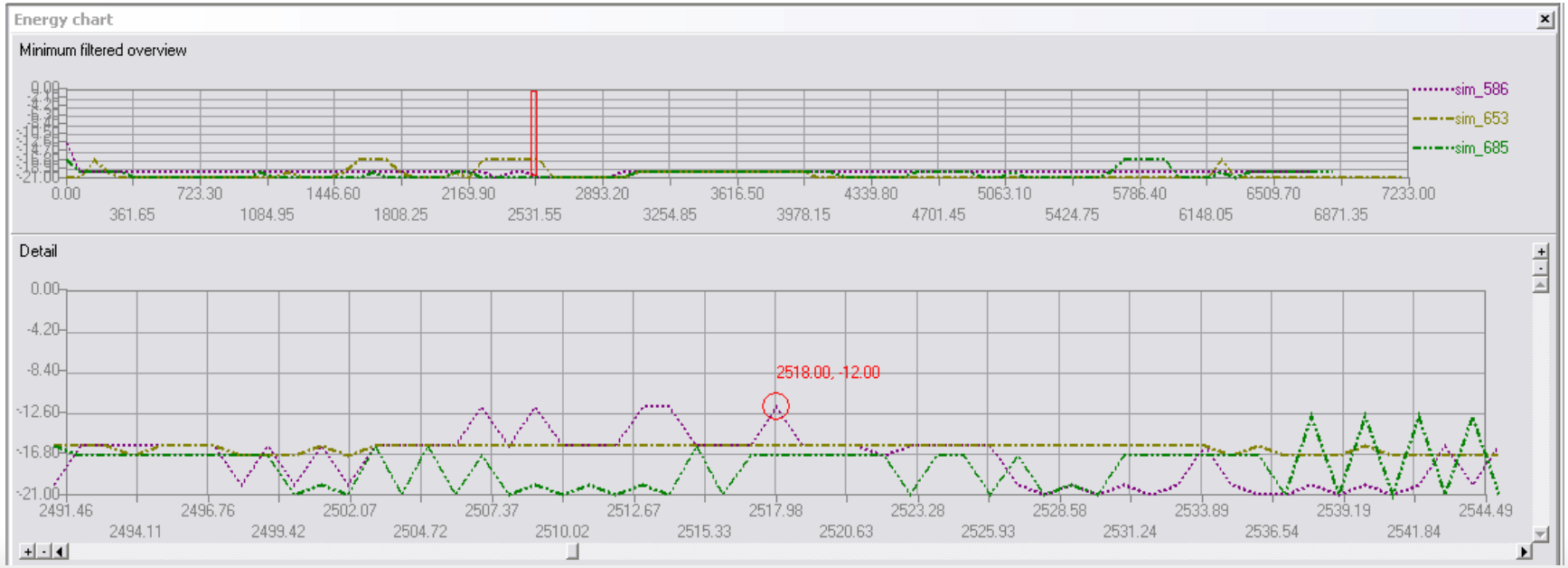
2.4 Detail views

“Overview first, zoom and filter, details on demand.”



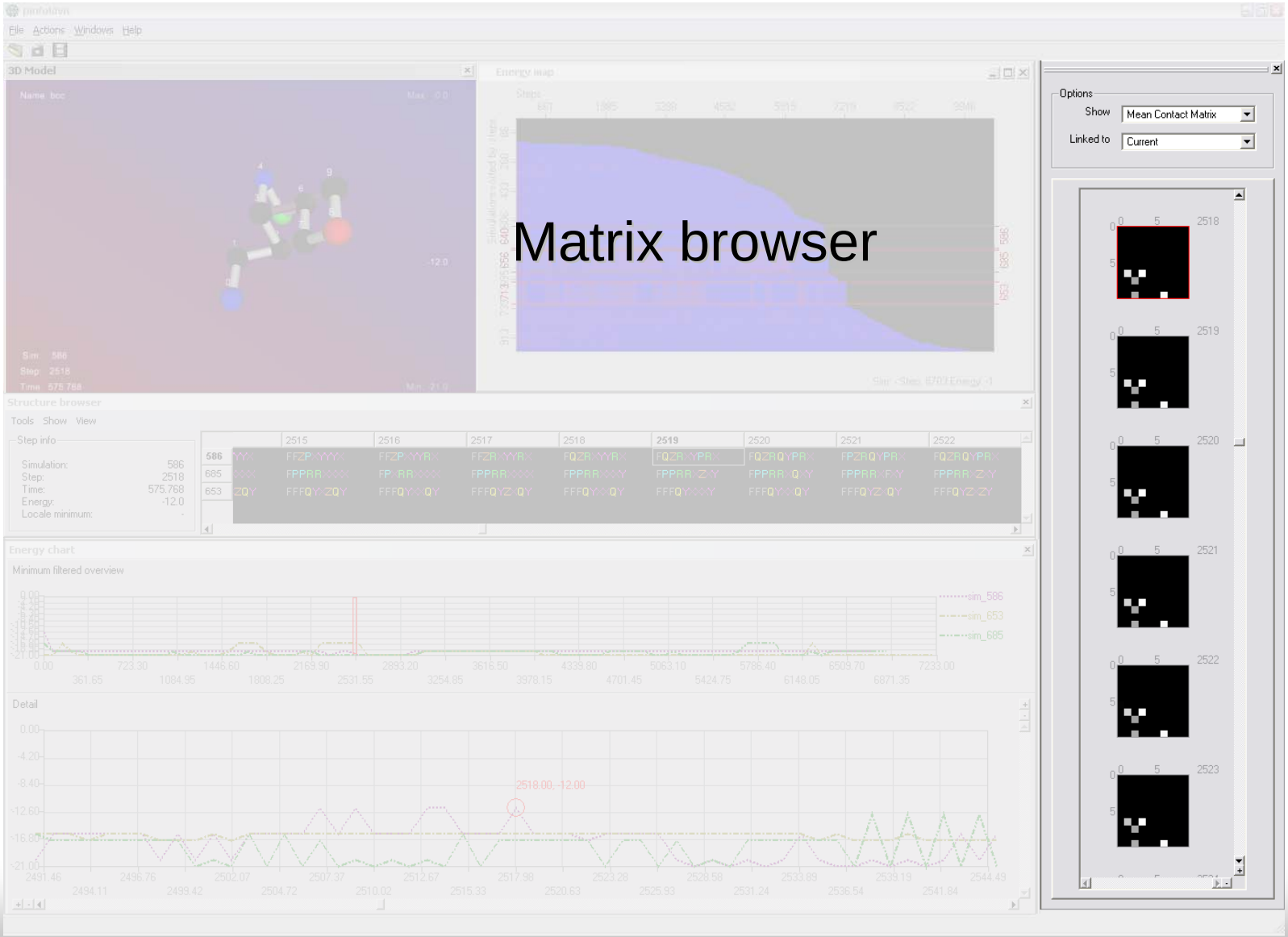
2.4 Detail views

- Energy viewer
- Overview+detail technique
  - Preserve the overview while exploring details
  - Space multiplexing or time multiplexing
  - Geometric zoom or semantic zoom



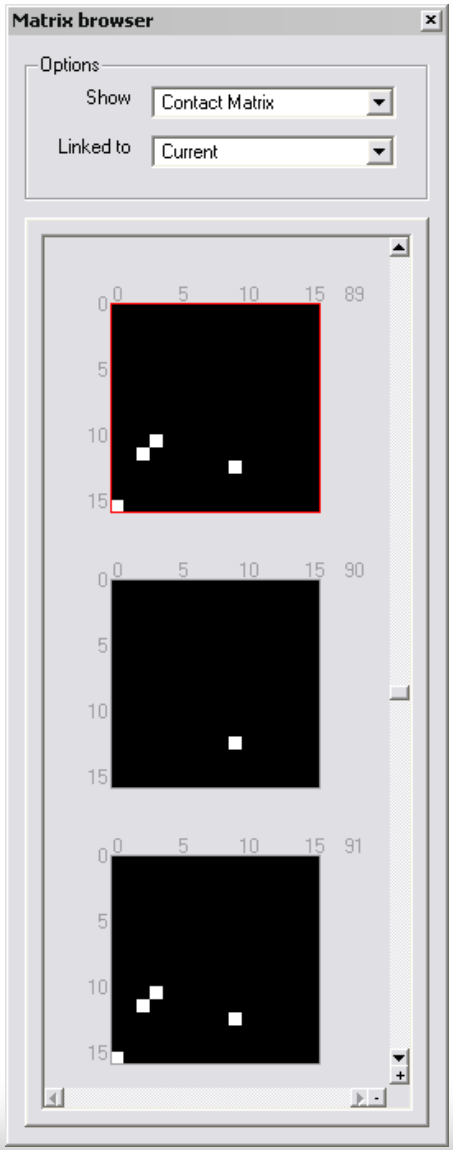
2.4 Detail views

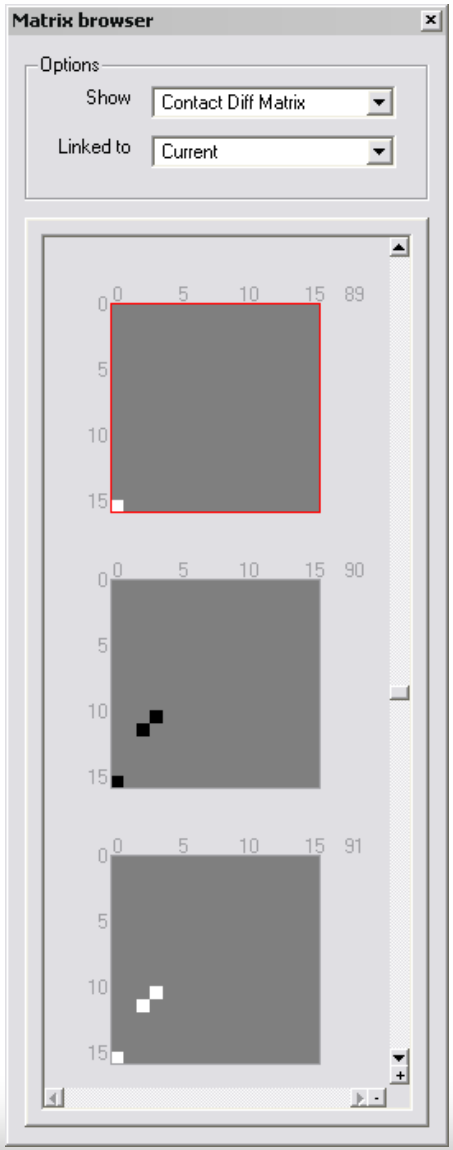
"Overview first, zoom and filter, details on demand."

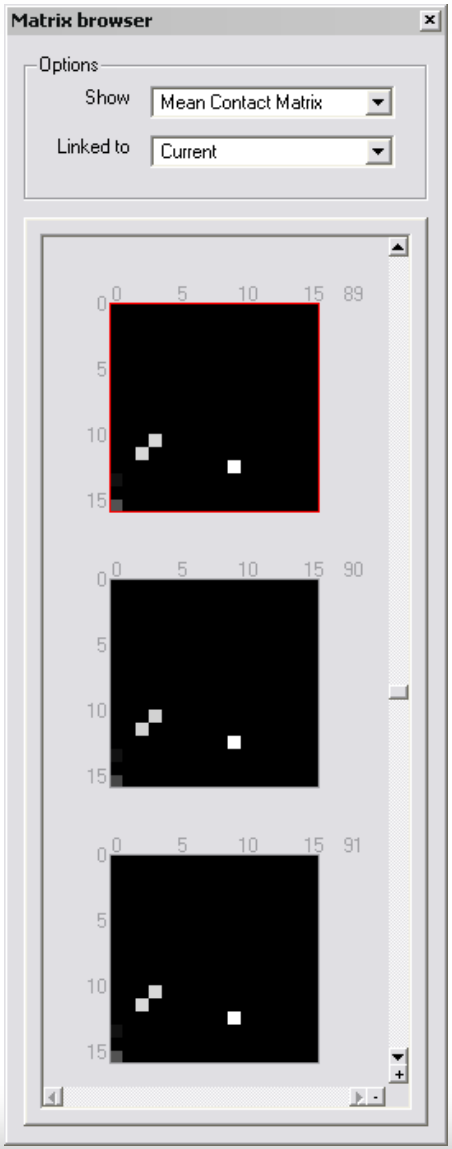


2.4 Detail views

- Matrix browser

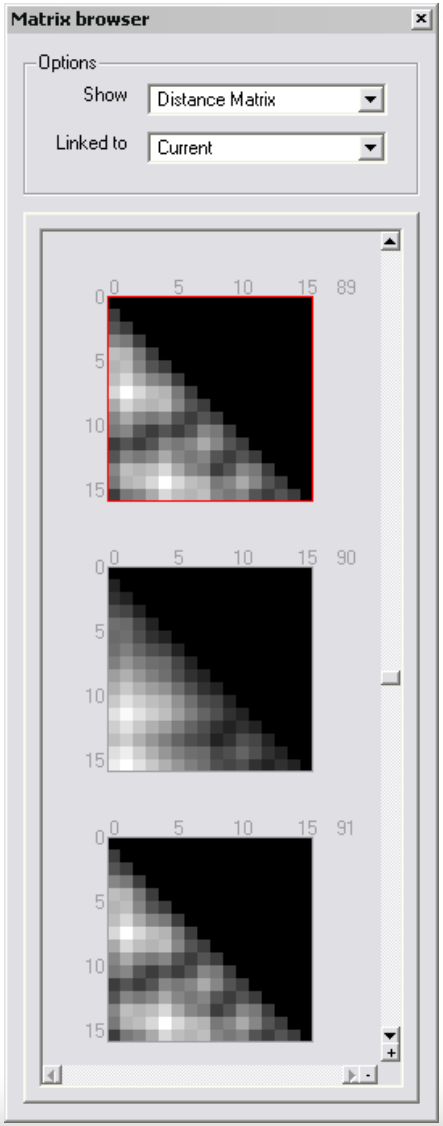






2.4 Detail views

- Matrix browser



3. Pinfoldvis Demo

- Tool demonstration

4. Bibliography

- *Baldonado, M. Q. W., Woodruff, A., and Kuchinsky, A. Guidelines for using multiple views in information visualization. In Advanced Visual Interfaces (2000), pp. 110–119.*
- *Dill, K. A. Polymer principles and protein folding. Protein Science 8 (1999), 1166–1180.*
- *Card, S. K., Mackinlay, J. D., and Shneiderman, B., Eds. Readings in Information Visualization — Using Vision to Think. Morgan Kaufmann, 1999.*
- *Will, S. Exact, constraint-based structure prediction in simple protein models. PhD thesis, Universität Jena, 2005.*
- *Wolfinger, M. Energy Landscapes of Biopolymers. PhD thesis, Fakultät für Chemie, Universität Wien, 2004.*