

# Scientific writing

**Ivan Viola**  
& collective



UNIVERSITETET I BERGEN

# Forms of Information Dissemination

- Scientific Paper
  - Journal publication
  - Conference publication (Proceedings)
- Poster
- Thesis (Master, PhD)
- Tutorial/Course Notes
- STAR-Report
- Invited talk
- Popular reports in news and broadcasting
- X-Mas Card?



# Important Journals and Conferences

## ■ Journals

- IEEE Trans. on Visualization and Computer Graphics (TVCG)
- IEEE Computer Graphics and Applications (CG&A)
- Transactions on Graphics (TOG)
- Computer Graphics Forum (CGF)

## ■ Conferences and Symposia

- IEEE Visualization
- Eurographics/IEEE VGTC Symposium on Visualization (EuroVis)
- ACM SIGGRAPH
- Eurographics

# Forms of Information Dissemination

## Scientific Paper

- Journal publication

- Conference publication (Proceedings)



- Poster

- Thesis (Master, PhD)

- Tutorial/Course Notes

- STAR-Report

- Invited talk

- Popular reports in news and broadcasting

- X-Mas Card?



# Scientific Publication

- Mostly used form of dissemination of novel ideas
- Target audience: specific scientific community
- Ensuring the quality: peer-reviewing, international program committee (IPC), program chairs
- Must read as a story
  - **Problem** (no overview on fish stocks, uncontrolled fishing)
  - **Goal** (sustainable fishery)
  - **Approach** (sonar scanning of sea regions and interactive visual analysis of data)
  - **Outcome** (precise quota estimation)
  - **Costs** (research, surveillance devices, manpower, ships)

# Paper Quality Evaluation

## ■ Scientific value

- Does it address an important problem?
- Level of novelty (significant step or incremental?)
- Technical soundness
- In scope of the journal/conference?

## ■ Presentation quality

- Clarity of presentation
- Problem description
- References the most important work in the field
- Reproducibility

# Scientific Publication

- Often describes difficult and complex approach
- Reader should be considered!

*If the reader is to grasp what the writer means, the writer must understand what the reader needs.*

*[G. Gopen and J. Swan 1990]*

1.  $t(\text{time})=15'$ ,  $T(\text{temperature})=32^\circ$ ;  $t=0'$ ,  $T=25^\circ$ ;  $t=6'$ ,  $T=29^\circ$ ;  $t=3'$ ,  $T=27^\circ$ ;  $t=12'$ ,  $T=32^\circ$ ;  $t=9'$ ;  $T=31^\circ$

- 2.

time (min)	temperature(°C)
0	25
3	27
6	29
9	31
12	32
15	32

- 3.

temperature(°C)	time (min)
25	0
27	3
29	6
31	9
32	12
32	15

# Style Guidelines

- Always keep the **red thread**!
- Be consistent among terms (set-up, setup, set up)
- Reasoning first, approach details afterwards
- Avoid using contractions (e.g., it's, can't, ... )
- No paragraph should have only one sentence
- Use a spellchecker
- Sentence should not be too long (max 2 lines)
- Use simple formulations and short sentences, if a sentence can be broken into two, do it.
- Try to avoid abbreviations (except: i.e., e.g.), if more convenient, introduce them first
- Avoid colloquiality



# Style Guidelines (cont.)

- Serif font is designed for easier reading the line
- Use sans-serif in titles but not in the plain text
- Include methodological illustrations, rendered images
- Equations variables have to be described in the text
- Equations, Figures, Tables all has to be referenced and discussed in the text labelled and having caption
- Captions - little redundancy is welcome
- Very important to explain first the high-level information and then to dig into the details
- Use consistency among terms
- Prefer positive formulations over negative ones

# Paper Structure

- Title
- Authors
- Abstract
- Keywords, Index Terms

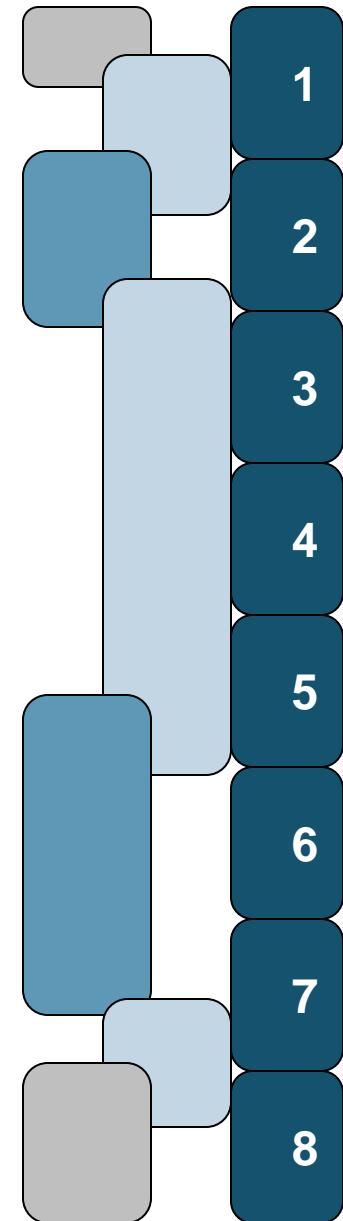
## Head

- Introduction
- Related Work
- Basic Idea
- Approach in Detail
- Experiments, Validation,
- Results, Performance, Discussion

## Body

- Outlook
- Conclusions
- References, Appendix

## Tail



# Illustrative Example

## Two-Level Approach to Efficient Visualization of Protein Dynamics

Ove Daas Lampe, Ivan Viola Member, IEEE Computer Society, Nathalie Reuter and Helwig Hauser Member, IEEE

**Abstract**—Proteins are highly flexible and large amplitude deformations of their structure, also called slow dynamics, are often decisive to their function. We present a two-level rendering approach that enables visualization of slow dynamics of large protein assemblies. Our approach is aligned with a hierarchical model of large scale molecular motions. Instead of constantly updating positions of large amounts of atoms, we update the position and rotation of molecules (i.e. higher level building blocks of a protein). Molecules are represented hierarchically on the GPU, exploiting the new graphics hardware geometry shader capabilities. Moreover, we represent the atoms by billboards instead of tessellated spheres. Our representation is then significantly faster and just precise. We demonstrate the usefulness of our new approach in the context of our collaborative bioinformatics project.

**Index Terms**—Molecular visualization, hardware acceleration, protein dynamics.

### 1 INTRODUCTION

Proteins are often considered to be static objects. However, proteins are highly dynamic, and their dynamics are often the key to their function [2]. For example, some proteins have an open and a closed form and understanding the transition between both is crucial to be able to design efficient drugs [17, 23]. Of particular interest are the large amplitude conformational changes (MAC) of proteins. Molecular dynamics (MD) are widely used to study the conformational changes of protein structures along time, but very long MD simulations of molecular systems containing hundreds of thousands of atoms are extremely demanding and appear to be computationally intractable. Normal mode analysis (NMA) with coarse grained atoms is much less computationally demanding and appears to be a better approach, as has been successfully applied to predict the conformational changes of, for example, nanobodies or virus capsids [13].

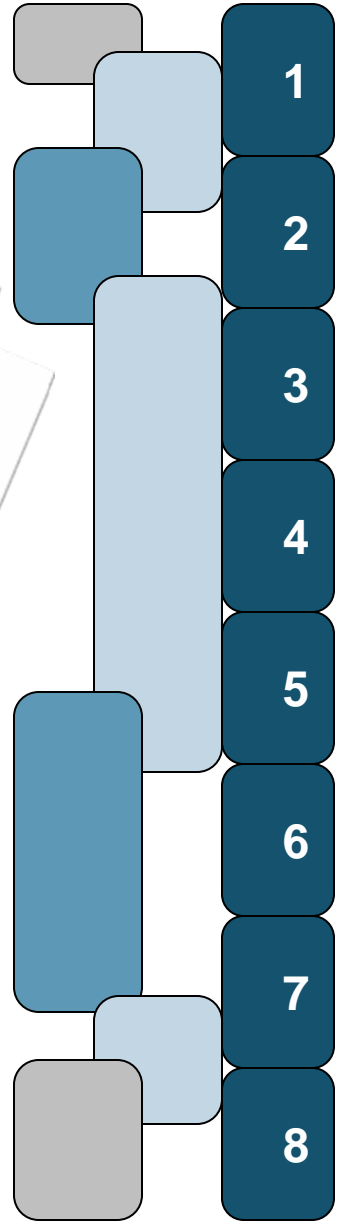
Proteins are made of amino acids, referred to as residues. In this work, we exploit the fact that protein structures can be divided into two hierarchical separate levels. Although there are only 20 different standard amino acids, which contain between 1 and 24 atoms, including the hydrogen atoms [16], some proteins need other interactions to stabilize their structure, such as water or other more complex molecules. To support such constructs, we use a model that allows mixing of two-level structures, like proteins and one-level structures such as water. All amino acids have a common structure (forming the so-called backbone) and they differ by side chains of different lengths and complexity. Refer to Fig. 1 for more detail with respect to the structure of proteins. Amino acids are linked by peptide bonds to form a chain. These chains often contain several bends or kinks, forming a complex structure. Several proteins can interact with each other and aggregate to form a complex assembly whose function can differ from the individual proteins. In fact, the biological structures we have to look at, for a better understanding of crucial biological processes, often contain a large number of atoms. With all the progress made in our post-genomic era, the number of atoms in biological structures under investigation steadily increases.

- Ove Daas Lampe is with Christian Michelsen Research, Norway. E-mail: ovel@cmr.no
  - Ivan Viola is with Department of Informatics, University of Bergen, Norway. E-mail: ivan.viola@iuh.uib.no
  - Nathalie Reuter is with Computational Biology Chair/CCO, University of Bergen, Norway. E-mail: nreuter@iuh.uib.no
  - Helwig Hauser is with Department of Informatics, University of Bergen, Norway. E-mail: helwig.hauser@iuh.uib.no
- Manuscript received 11 March 2007; accepted 1 August 2007; posted online 27 October 2007.  
For information on obtaining reprints of this article, please contact the publisher.

One way of gaining insight into the dynamics of proteins is to analyze the protein structure using two-dimensional plots or molecular tables. A more intuitive approach is also making a 3D visualization. Interactive 3D visualization of protein dynamics, for example, has been demonstrated to serve as a very good exchange platform between experimental and computational biologists. Experimental biologists are not familiar with the methods employed by computational biologists to represent and connect them with the protein structure and conformational changes. Interactive 3D visualization is more intuitive and enables the visualization of protein dynamics in a much more understandable way. The protein structure is also made of a much larger number of atoms than the 2D visualization. A more intuitive way of depicting the protein structure is to use a hierarchical representation of the protein structure. A more intuitive way of depicting the protein structure is to use a hierarchical representation of the protein structure. A more intuitive way of depicting the protein structure is to use a hierarchical representation of the protein structure.

There are several different visual representations in use for the visualization of protein structures. Each visual representation serves a specific purpose. Some of them depend on high-level information given as part of the backbone. Such a visual depiction is referred to as a coarse-grained representation. Some of them depend on the atom in the protein with the highest resolution. Such a visual depiction is referred to as a fine-grained representation. Some of them depend on the atom in the protein with the highest resolution. Such a visual depiction is referred to as a fine-grained representation. Some of them depend on the atom in the protein with the highest resolution. Such a visual depiction is referred to as a fine-grained representation.

Interactive visualization of very large proteins (hundreds of thousands of atoms) using the latest graphics hardware is computationally very demanding, even using the latest graphics hardware. In our project, this task goes even more challenging since protein dynamics, especially at a high resolution, is associated with dynamic geometry changes of the protein representation. Finding new geometry to represent the atoms is a non-trivial task. Especially in the case of an atom, the geometry changes are often very large. In our project, we use a two-level approach to visualize protein dynamics. The basic idea builds upon the natural hierarchical structure of proteins. The basic idea builds upon the natural hierarchical structure of proteins.





# Illustrative Example

## Two-Level Approach to Efficient Visualization of Protein Dynamics

Ove Daae Lampe, Ivan Viola Member, IEEE Computer Society, Nathalia Reuter and Helwig Hauser Member, IEEE

**Abstract**—Proteins are highly flexible and large amplitude deformations of their structure, also called slow dynamics, are often decisive to their function. We present a two-level rendering approach that enables visualization of slow dynamics of large protein assemblies. Our approach is aligned with a hierarchical model of large scale molecular motions. Instead of constantly updating positions of large amounts of atoms, we update the position and rotation of molecular subunits (i.e. higher-level building blocks of a protein). Subunits are represented by one vertex only reducing the number and additional information defining the model. The atoms are represented as spheres by billboards instead of textured spheres. Our representation is then significantly faster and just precise. We demonstrate the usefulness of our new approach in the context of our collaborative bioinformatics project.

**Index Terms**—Molecular visualization, hardware acceleration, protein dynamics.

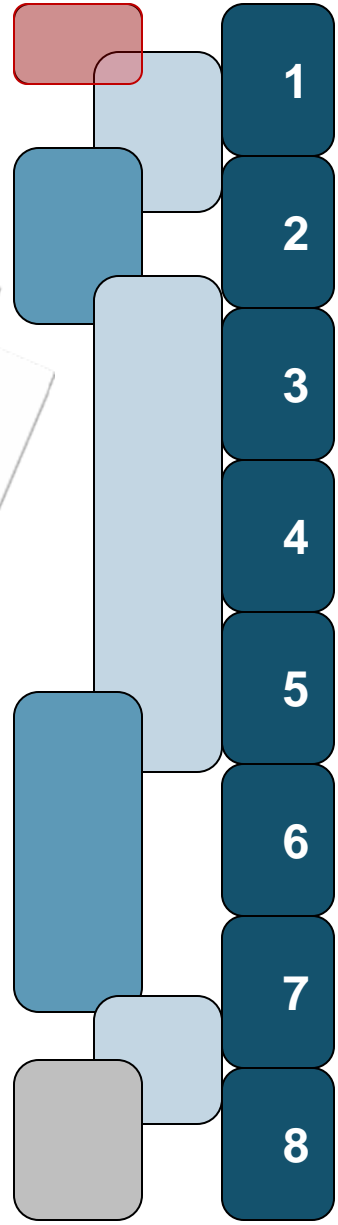
### 1 INTRODUCTION

Proteins are often considered to be static objects. However, proteins are highly dynamic, and their dynamics are often the key to their function [2]. For example, some proteins have an open and a closed state and understanding the transition between both is crucial to be able to design efficient drugs [17, 23]. Of particular interest are the large amplitude conformational changes (MOC) which are widely used to study the conformational changes of protein structures along time, but very long MD simulations of molecular systems containing hundreds of thousands of atoms are often very challenging. Normal mode analysis (NMA) with coarse grained models is much less computationally demanding and appears to be a better approach, it has been successfully applied to predict the conformational changes of, for example, nanoprobes or virus capsids [13].

Proteins are made of amino acids, referred to as residues. In this work, we exploit the fact that protein structures can be divided into two hierarchical separate levels. Although there are only 20 different standard amino acids, which contain between 1 and 24 atoms, also including the hydrogen atoms [16], some proteins need other reference constructs to make sense, such as water or other more complex molecules. To support such constructs, we need a model that allows mixing of two-level structures, like proteins and one-level structures such as water. All amino acids have a atoms in common (forming the backbone) and they differ by side chains of different length and complexity. Refer to Fig. 1 for more detail with respect to the structure of proteins. Amino acids are linked by peptide bonds to form a chain. These chains often contain several hundreds of amino acids. A protein can be made of one or several chains. In living organisms, several proteins can interact with each other and aggregate to form a complex assembly whose function can differ from the individual parts. In fact, the biological structures we have to look at, for a better understanding of crucial biological processes, often contain a large number of atoms. With all the progress made in our post-genomic era, the number of atoms in biological structures under investigation steadily

- Ove Daae Lampe is with Christian Michelsen Research, Norway. E-mail: ovd@cmr.no
  - Ivan Viola is with Department of Informatics, University of Bergen, Norway. E-mail: ivan.viola@iuh.uib.no
  - Nathalia Reuter is with Computational Biology Unit/CCO, University of Bergen, Norway. E-mail: nathalia.reuter@iuh.uib.no
  - Helwig Hauser is with Department of Informatics, University of Bergen, Norway. E-mail: helwig.hauser@iuh.uib.no
- Manuscript received 11 March 2007; accepted 1 August 2007; revised October 2007.  
 For information on obtaining reprints of this article, please contact the publisher.

One way of gaining insight into the dynamics of proteins is to analyze the protein structure using two-dimensional plots or molecular motion visualization. The protein structure is also modeled as a much more understandable representation of the protein structure. For example, the protein structure is represented as a very good example of a hierarchical model of large scale molecular motions. Instead of constantly updating positions of large amounts of atoms, we update the position and rotation of molecular subunits (i.e. higher-level building blocks of a protein). Subunits are represented by one vertex only reducing the number and additional information defining the model. The atoms are represented as spheres by billboards instead of textured spheres. Our representation is then significantly faster and just precise. We demonstrate the usefulness of our new approach in the context of our collaborative bioinformatics project.



- Thousands of people read the title without reading rest!
- Quoting Master: "Kurz, prägnant, arrogant!"  
(*short, concise, arrogant*)
- Two (extreme) examples from IEEE Vis conference
  - ***Superellipsoid-based, Real Symmetric Traceless Tensor Graphs Motivated by Nematic Liquid Crystal Alignment Visualization***
    - Describes exactly what it is about but distracts a reader right at the beginning

uhhh brrrrr, wwwwwwhat?
  - ***Caricaturistic Visualization***
    - Too general title, hard to find out what the paper specifically address

hmm, sounds interesting...

■ Correct title design lies somewhere in between



# Illustrative Example

## Two-Level Approach to Efficient Visualization of Protein Dynamics

Ove Daae Lampe, Ivan Viola Member, IEEE Computer Society, Nuthaba Rauter and Helwig Hauser Member, IEEE

**Abstract**—Proteins are highly flexible and large amplitude deformations of their structure, also called slow dynamics, are often decisive to their function. We present a two-level rendering approach that enables visualization of slow dynamics of large protein assemblies. Our approach is aligned with a hierarchical model of large scale molecular motion. Instead of constantly updating positions of large amounts of atoms, we update the position and rotation of molecular blocks (i.e. higher-level building blocks of a protein). Positions are represented by atoms, only including its position and additional information defining the model. The atoms are represented as spheres by spheres instead of textured spheres. Our representation is then significantly faster and just precise. We demonstrate the usefulness of our new approach in the context of our collaborative bioinformatics project.

**Index Terms**—Molecular visualization, hardware acceleration, protein dynamics.

### 1 INTRODUCTION

Proteins are often considered to be static objects. However, proteins are highly dynamic, and their dynamics are often the key to their function [2]. For example, some proteins have an open and a closed state and are able to design efficient drugs [17, 23]. Of particular interest are the large amplitude conformational changes (MD) are widely used to study the conformational changes of protein structures along time, but very long MD simulations of molecular systems containing hundreds of thousands of atoms remain challenging. Normal modes analysis (NMA) with coarse grained models is much less computationally demanding and appears to be a better approach, as has been successfully applied to predict the conformational changes of, for example, nanobodies or virus capsids [13].

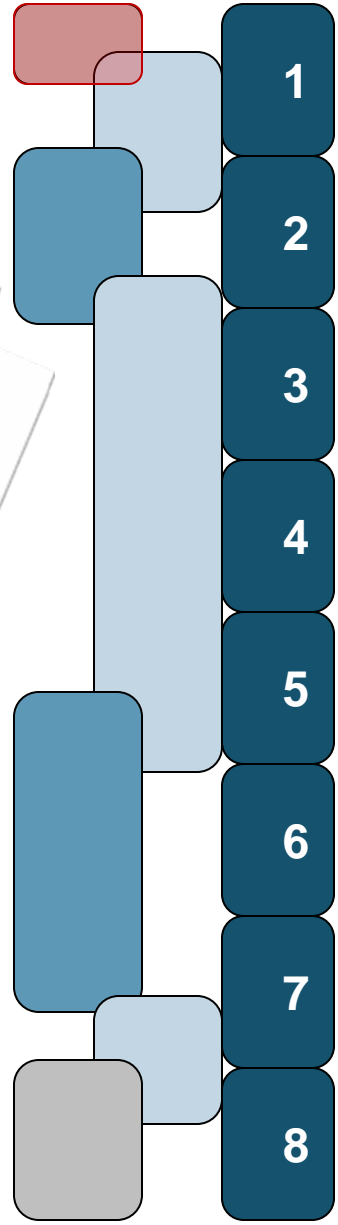
Proteins are made of amino acids, referred to as residues. In this work, we exploit the fact that protein structures can be divided into two hierarchical separate levels. Although there are only 20 different standard amino acids, which contain between 1 and 24 atoms, also including the hydrogen atoms [16], some proteins need other interactions to support such a structure, such as water or other more complex molecules. To support such a structure, we need a model that allows modeling of two-level structures, like proteins and one-level structures such as water. All amino acids have a atoms in common (forming the protein backbone) and they differ by side chains of different length and complexity. Refer to Fig. 1 for more detail with respect to the structure of proteins. Amino acids are linked by peptide bonds to form a chain. These chains often contain several hundreds of amino acids. A protein can be made of one or several chains. In living organisms, several proteins can interact with each other and aggregate to form a complex assembly whose function can differ from the individual components. In fact, the biological structures we have to look at, for a better understanding of crucial biological processes, often contain a large number of atoms. With all the progress made in our post-genomic era, the number of atoms in biological structures under investigation steadily

increases. One way of gaining insight into the dynamics of proteins is to analyze the protein structure using two-dimensional plots or molecular motion visualization. The results using two-dimensional plots or molecular motion visualization are often used to analyze the protein structure. For example, the visualization of protein dynamics using two-dimensional plots or molecular motion visualization is a very good exchange platform between experimental and computational biologists. Experimental biologists are not familiar with the methods employed by computational biologists (e.g. NMA [13]), which makes it difficult for them to analyze computational results and connect them with the protein structure and function. Interactive 3D visualization is more intuitive and results in a better depiction of protein structures. Each visual representation serves a different purpose. Some of them depend on the level of information given (e.g. the backbone, the side chains, etc.). Some rendering using the decrease resolution is referred to as coarse-grained rendering of the atoms (CG) and hydrogen atoms are not shown. In this way of depicting the protein structure, the atoms are not shown. In the coarse-grained rendering, the atoms are not shown. In the coarse-grained rendering, the atoms are not shown. In the coarse-grained rendering, the atoms are not shown.

There are several different visual representations in use for the visualization of protein structures. Each visual representation serves a different purpose. Some of them depend on the level of information given (e.g. the backbone, the side chains, etc.). Some rendering using the decrease resolution is referred to as coarse-grained rendering of the atoms (CG) and hydrogen atoms are not shown. In this way of depicting the protein structure, the atoms are not shown. In the coarse-grained rendering, the atoms are not shown. In the coarse-grained rendering, the atoms are not shown.

- Ove Daae Lampe is with Department of Informatics, University of Bergen, Norway. E-mail: ove.daae@iuh.uib.no
- Ivan Viola is with Department of Informatics, University of Bergen, Norway. E-mail: ivan.viola@iuh.uib.no
- Nuthaba Rauter is with Department of Informatics, University of Bergen, Norway. E-mail: nuthaba.rauter@iuh.uib.no
- Helwig Hauser is with Department of Informatics, University of Bergen, Norway. E-mail: helwig.hauser@iuh.uib.no

Manuscript received 21 March 2012; revised 12 July 2012; accepted 12 October 2012.  
 21 October 2012  
 For information on obtaining reprints of this article, please contact our reprints department.



This work was supported by the Norwegian Research Council (RCN) through the project 'Efficient Visualization of Protein Dynamics' (VITA) under grant number 169449. The authors would like to thank the anonymous reviewers for their constructive comments. We also thank the reviewers for their constructive comments.



# Authors and Affiliation

- Often several authors contribute to one publication
- First author: main responsible, main research contributor, work coordinator, trigger
  - Often PhD student, ...
- Co-authors → *et al.* (et alii, and others)
  - Project manager (PostDoc?)
  - Students
  - Supervisor (usually the last author)
- Affiliation – research institution
  - Several affiliations per person are allowed
  - Authors connected to affiliations through markups



# Illustrative Example

## Two-Level Approach to Efficient Visualization of Protein Dynamics

Ove Daas Lampe, Ivan Viola Member, IEEE Computer Society, Nathalia Reuter and Helwig Hauser Member, IEEE

**Abstract.** Proteins are highly flexible and large amplitude deformations of their structure, also called their dynamics, are often decisive to their function. We present a two-level rendering approach that enables visualization of slow dynamics of large protein assemblies. Our approach is aligned with a hierarchical model of large scale molecules. Instead of constantly updating positions of large amounts of atoms, we update the position and rotation of molecules on a higher level looking blocks of a protein. Blocks are represented by one vertex only reducing the number of vertices and triangles during the rendering. The atoms are represented by spheres by default instead of textured spheres. Our representation is then significantly faster and just previous, we demonstrate the usefulness of our new approach in the context of our collaborative bioinformatics project.

### 1 INTRODUCTION

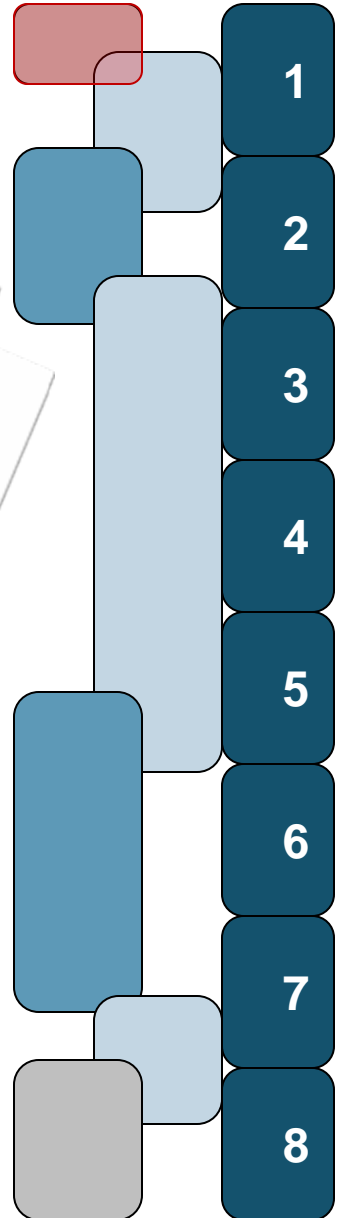
Proteins are often considered to be static objects. However, proteins are highly dynamic, and their dynamics are often the key to their function [2]. For example, some proteins have an open and a closed form and understanding the transition between both is crucial to be able to design efficient drugs [17, 23]. Of particular interest are the large amplitude conformational changes (MD) are widely used to study the conformational changes of protein structures along time, but very long MD simulations of molecular systems containing hundreds of thousands of atoms remain challenging. Normal mode analysis (NMA) with coarse grained models is much less computationally demanding and appears to be a better approach, it has been successfully applied to predict the collective large amplitude motions of, for example, nanoprobes or virus capsids [13].

Proteins are made of amino acids, referred to as residues. In this work, we exploit the fact that protein structures can be divided into two hierarchical separate levels. Although there are only 20 different standard amino acids, which contain between 1 and 24 atoms, also including the hydrogen atoms [16], some proteins need other reference constructs to make sense, such as water or other more complex molecules. To support such constructs, we used a model that allows nesting of two-level structures, like proteins and one level structures such as water. All amino acids have a atoms in common (forming the protein backbone) and they differ by side chains with respect to the size and complexity. Refer to Fig. 6 for more detail with respect to the protein structure. Amino acids are linked by peptide bonds to form a linear chain. These chains often contain several blocks of amino acids. A protein can be made of one or several chains. In living organisms, several proteins can interact with each other and aggregate to form a complex assembly whose function can differ from the individual parts. In fact, the biological structures we have to look at, for a better understanding of crucial biological processes, often contain a large number of atoms. With all the progresses made in our post-genomic era, the number of atoms in biological structures under investigation steadily

- Ove Daas Lampe is with Christian Michelsen Research, Norway. E-mail: ove@lampe.com
  - Ivan Viola is with Department of Informatics, University of Bergen, Norway. E-mail: ivan.viola@iuh.uib.no
  - Nathalia Reuter is with Computational Biology Center (CBCC), University of Bergen, Norway. E-mail: nathalia.reuter@iuh.uib.no
  - Helwig Hauser is with Department of Informatics, University of Bergen, Norway. E-mail: helwig.hauser@iuh.uib.no
- Manuscript received 11 March 2007; accepted 7 August 2007; revised online 27 October 2007.  
For information on obtaining reprints of this article, please contact our publisher.

One way of gaining insights into the dynamics of proteins is to analyze the geometric information arising from the NMA calculations and visualize the results using two-dimensional plots or scatter plots. A more intuitive approach which also enables a much better understanding of the spatial relationships in dynamic protein dynamics, for example, the potential and computational biologists, experimental biologists are often not familiar with the methods employed by computational biologists in e.g., NMA [13], which makes it difficult for them to analyze computational results and connect them with the protein structure and function. Interactive 3D visualization is more intuitive and enables in-depth depiction of protein structures. Each visual representation serves a specific purpose. Some of them depend on high level information giving some rendering using the coarse structure and investigating how a part of the backbone, such a visual depiction is referred to as coarse-grained representation of the protein structure. A more traditional but still often preferred way of depicting the structure is known as ball-and-stick representation where each atom in the structure is shown as ball and they interact with each other. Smoothing spheres are often used to represent the atoms and the interaction between them is often represented by connecting lines (see Fig. 1). Extract structural information from a 3D representation, additional techniques like depth sorting, and complex structures [27] are vital to extract structural information from a 3D representation of a protein structure. In our project, we do not lose any detail, but we use a hierarchical representation of the protein structure. The complete geometry of the protein representation is associated with dynamic geometry information using the latest graphics hardware. In our project, this task gives even more challenging since protein dynamics are often simulated on a high resolution projection with dynamic geometry information. Simulating protein dynamics on a high resolution projection of the protein representation is associated with dynamic geometry information. Simulating protein dynamics on a high resolution projection of the protein representation is associated with dynamic geometry information. Simulating protein dynamics on a high resolution projection of the protein representation is associated with dynamic geometry information.

To achieve interactive rendering in the above mentioned scenario, we use a coarse two-level approach for rendering dynamic protein visualization technology and interaction. To achieve interactive rendering in the above mentioned scenario, we use a coarse two-level approach for rendering dynamic protein visualization technology and interaction. To achieve interactive rendering in the above mentioned scenario, we use a coarse two-level approach for rendering dynamic protein visualization technology and interaction.





# Abstract and Keywords

- Abstract includes
  - Questions you investigated
  - One sentence motivation for the work
  - Short description of the method
  - Advantages over existing methods
  - A brief summary of conclusions
- Ca 200 words not longer
- Motivate reader to read further
- Keywords
  - To enable easy indexing
  - Explicit indexing through classification system

<http://www.acm.org/class/1998/>

# Illustrative Example

## Two-Level Approach to Efficient Visualization of Protein Dynamics

Ove Daas Lampe, Ivan Viola Member, IEEE Computer Society, Nathalie Reuter and Helwig Hauser Member, IEEE

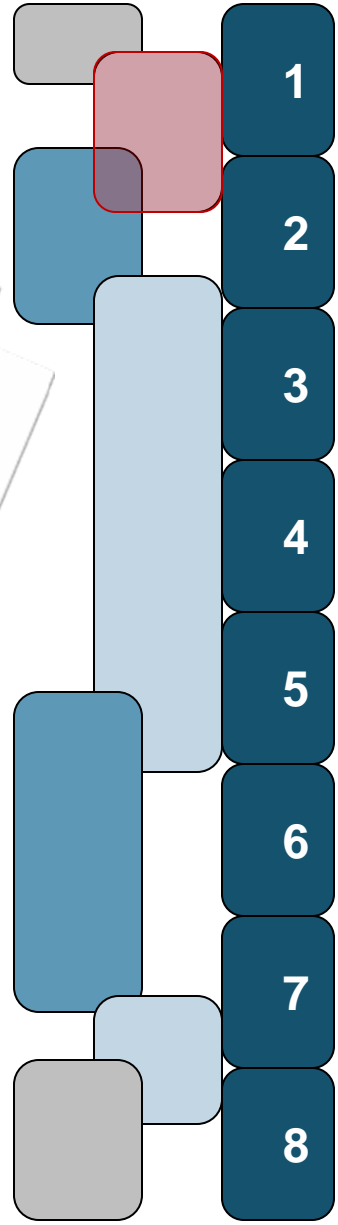
**Abstract**—Proteins are highly flexible and large amplitude deformations of their structure, also called slow dynamics, are often decisive to their function. We present a two-level rendering approach that enables visualization of slow dynamics of large protein assemblies. Our approach is aligned with a hierarchical model of large scale molecular motions. Instead of constantly updating positions of large amounts of atoms, we update the position and rotation of residues (i.e. higher level building blocks of a protein). Residues are represented by one vertex only reducing the number of vertices to be processed. The atoms are represented as spheres by tabulated instead of tessellated spheres. Our representation is then significantly faster and just precise. We demonstrate the usefulness of our new approach in the context of our collaborative bioinformatics project.

**Index Terms**—Molecular visualization, hardware acceleration, protein dynamics.

Proteins are often considered to be static objects. However, proteins are highly dynamic, and their dynamics are often the key to their function [2]. For example, some proteins have an open and a closed form and understanding the transition between both is crucial to be able to design efficient drugs [17, 23]. Of particular interest are large amplitude conformational changes (MOCs) which are often the result of functional changes of protein structures along time, but very slow MOCs are challenging. Normal mode analysis (NMA) with respect to the conformational changes of proteins containing hundreds of thousands of atoms is a well established method, but it is computationally expensive and a better approach is to use coarse-grained models. In this paper, we present a better approach to visualize MOCs, referred to as residues. In this paper, we present a better approach to visualize MOCs, referred to as residues. In this paper, we present a better approach to visualize MOCs, referred to as residues.

One way of gaining insight into the dynamics of proteins is to analyze the protein structure using two-dimensional plots or molecular motion analysis. This is done by projecting the protein structure onto a 2D plane. This is done by projecting the protein structure onto a 2D plane. This is done by projecting the protein structure onto a 2D plane.

- Ove Daas Lampe is with Department of Informatics, University of Bergen, Norway. E-mail: ove.lampe@iuh.uib.no
  - Ivan Viola is with Department of Informatics, University of Bergen, Norway. E-mail: ivan.viola@iuh.uib.no
  - Nathalie Reuter is with Department of Informatics, University of Bergen, Norway. E-mail: nathalie.reuter@iuh.uib.no
  - Helwig Hauser is with Department of Informatics, University of Bergen, Norway. E-mail: helwig.hauser@iuh.uib.no
- Manuscript received 11 March 2007; accepted 1 August 2007; revised 20 October 2007.  
For information on obtaining reprints of this article, please send e-mail to: reprints@computer.org.



# Introduction

- Background knowledge
  - in visualization often application domain specifics
- Problem description
- Stating objectives
- Main contribution of your approach
- An nice resulting rendered image in the front is always good as an appetizer  
(relevant mostly for Vis and CG)



# Related Work

- Put the work into context of other approaches
- Clearly state what has been known until now
- Cite every relevant work
- Provide brief comparison to your approach
- Citations writing style:
  - Two authors use Surname1 **and** Surname2
  - More authors use Surname1 et al. (plural!)
  - Avoid configurations like:
    - ... Dahl et al. [Dahl et al. 2007] ... (`\shortcite{}`)
  - By removing reference the sentence should still make sense

# Illustrative Example

means for mid-sized proteins (ca. 10,000 atoms) in a desktop environment through simplifying geometry that represents a single atom.

Reducing geometry complexity and increasing performance by depicting the details using image-based representations is a very promising technique in various fields of computer graphics [17]. Recently, billboard-based techniques have been applied for molecular visualization such as Chaperon [27]. Billboarding results in very high frame rates even for reasonably sized non-dynamic molecules. Our rendering of single atoms extends this billboarding technique. As Quaternions are limited to orthogonal projections only, we use the technique proposed by Gatzke [5] for rendering perspective-correct spheres, in the context of the ball-and-stick representation to support the dynamic visualization of large atomic structures for stress analysis [25, 23] to render the bonds.

Our two-level rendering approach is especially designed to support the dynamic visualization of large atomic structures for stress analysis. In addition to the perspective-correct spheres, we utilize focus-correct projection [7] for interactive exploration, we utilize focus-correct projection [7] for interactive exploration, we utilize focus-correct projection [7] for interactive exploration.

In this paper, we demonstrate how the novel graphics hardware capabilities are utilized to come up with an elegant two-level approach to interactive visualization of large and at the same time dynamic protein structures. Additionally, we improve and integrate the above reviewed techniques for billboarding, perspective correction, and repulsive interaction to better suit our purposes.

**INTERACTIVE DYNAMIC PROTEINS**

To understand the spatial relationships in the interaction amongst several dynamic proteins, biologists need to study such behavior using stereo projection and on high-resolution projection walls. To satisfy the scene complexity, our approach cannot be compromised by the extent of the scene complexity. In the following we describe a new rendering approach which enables the required interactivity, even for large protein structures. Position updates for every single atom in large protein structures in every single frame would normally stop the rendering process due to limited CPU bandwidth. By exploiting the fact that atoms are stationary in the NMA simulation only, we can avoid the need for calculating vectors for backbone elements and then afterwards apply them to all atoms, we can avoid the need for calculating vectors for backbone elements and then afterwards apply them to all atoms, we can avoid the need for calculating vectors for backbone elements and then afterwards apply them to all atoms.

In the first level of our new rendering pipeline, dynamic changes in the protein structure are represented by a control point, i.e., by a corresponding vertex in the graphics card. In the second rendering level we dynamically transfer the atoms which are contained in the residues, utilizing the geometry shader as featured in the latest graphics hardware generation [16]. For every atom we emit four vertices from the control point of the residue and represent the atoms in perspective-correct projection. The shading of the spheres which represent the atoms in the fragment shader is described in Figure 2 and the individual steps are described in detail in the following subsection.

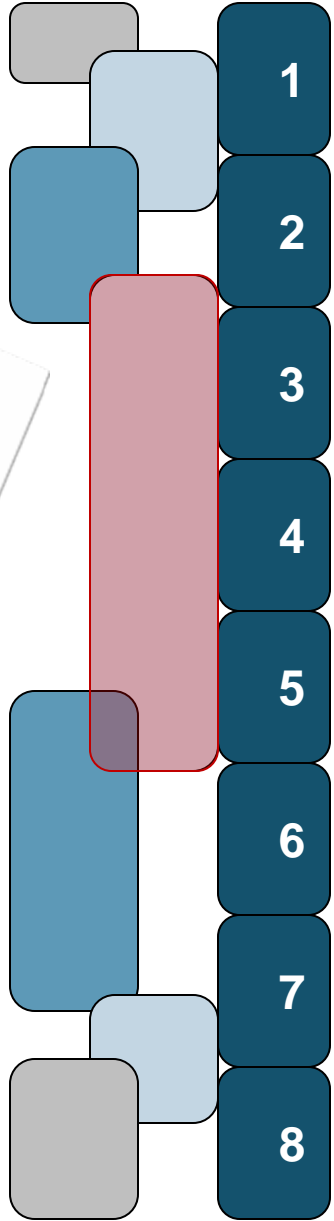
**5.4 Two-Level Rendering Pipeline**

Prior to rendering protein structures, we parse the input files [24] that define the protein and generate the protein hierarchy. All residues in the protein chain are identified and attached to the backbone atom which is present in every residue and which is denoted as the backbone control point. This atom serves as the origin of the local coordinate system of the residue. The atoms in the protein are assigned to respective residues and their relative position to the control point is computed. This transformation is defined as translation from the origin of the residue.

For each residue we store the following information:

- control point position  $(X, Y, Z)$  - residue position array
- rotations  $\alpha, \beta, \gamma$  - residue rotation array
- index of first atom  $A$  - atom offset array

Most recent approaches are focusing on displaying the dynamics of every atom in the molecule. Hao et al. [9] achieve interactive trans-

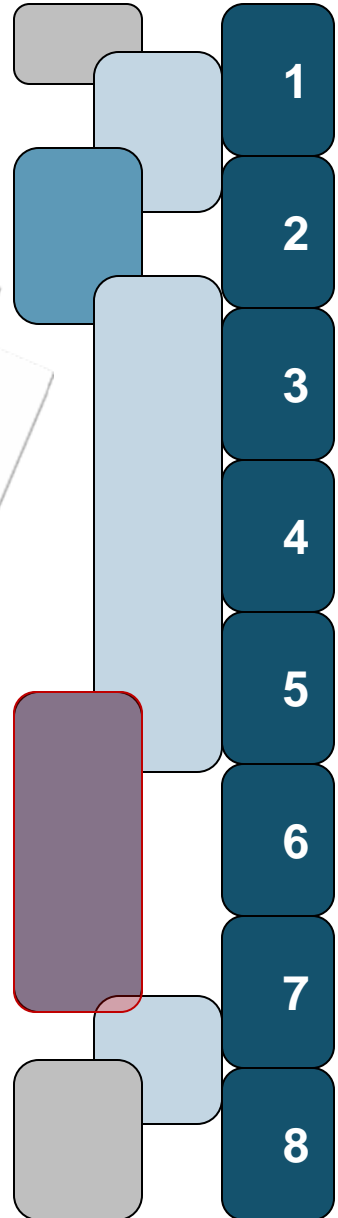
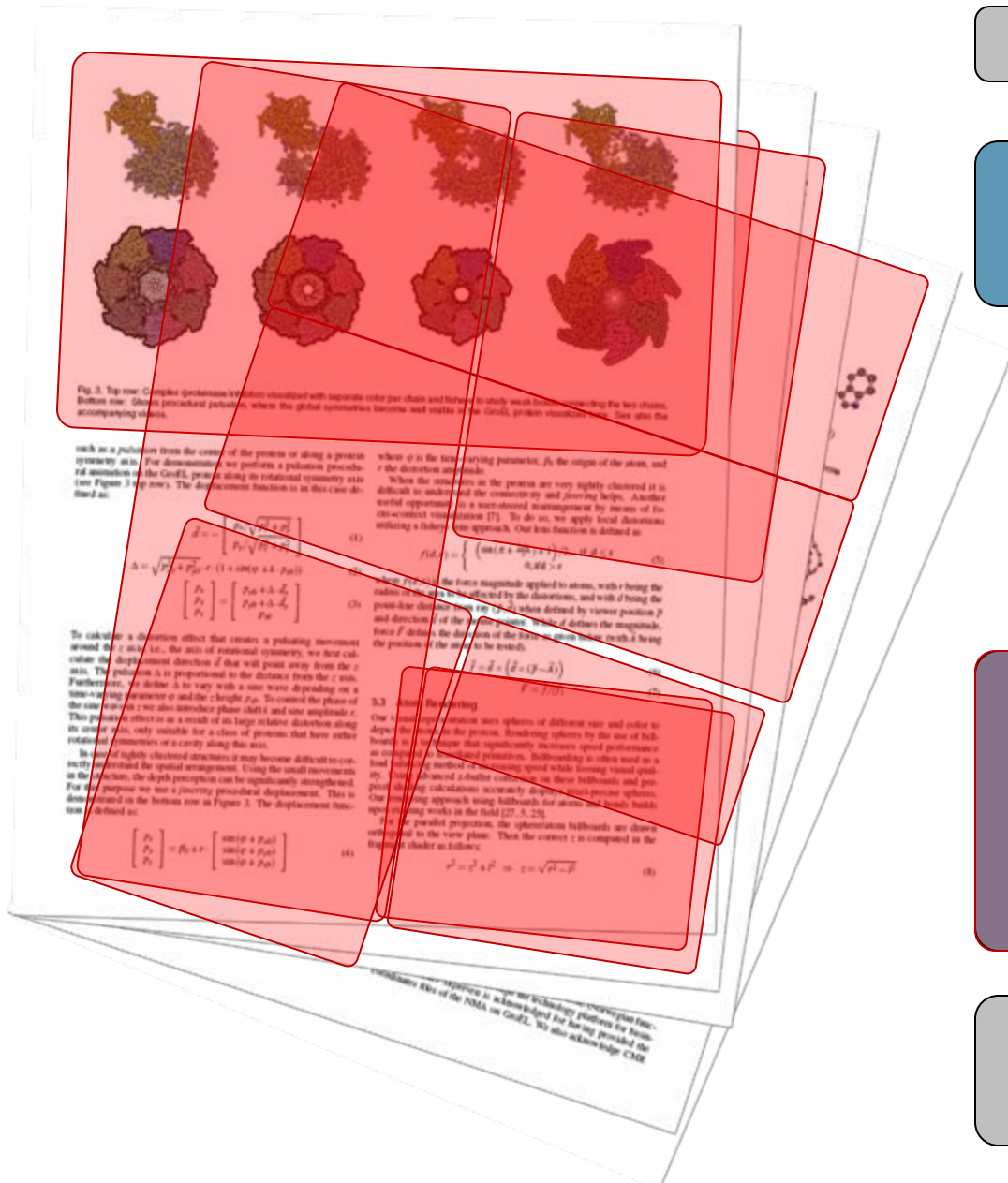


# Main Sections

- The contribution of your work has to be presented as simple as possible
  - Start with the big picture
  - A workflow, pipeline, or framework graphics is very useful - visual table of contents
  - Describe the algorithm in detail
  - Specific information on implementation
- Use illustrative graphics, tables, equations, pseudocode and discuss them in the main text
- Every statement should be either proven or referenced from past
- Reviewer has short time to evaluate, provide answers to her/his potential questions



# Illustrative Example





# Validation and Results

- Demonstration of usefulness in the application domain
- Resulting rendered images especially important
  - Figure placement should be *close* to discussion addressing the figure
  - Optimally on the same page or double-page
- Performance analysis
- Convergence behaviour
- Data description
- Error analysis

# Illustrative Example

web-based tool to calculate large amplitude movements of a protein, starting from the Cartesian coordinates of C $\alpha$  atoms [10]. This tool, based on the Molecular Modeling Tool Kit [9], is called WIZARD and performs Normal Modes Analysis (NMA) of proteins [8]. Among other things, this tool allows to create an animation of the movement of a protein in response to a trigger. The animation can be visualized within the web browser or downloaded and visualized with VMD. However, in both cases we display only the C $\alpha$  and still the hinges in Figure 5, on the other side, are frames from our new interactive visualization of the chequerboard GroEL dynamics.

The GroEL protein [11] contains 53,822 atoms and the animation not only the C $\alpha$  atoms, but also the side chains. This is an important level of information, but also the side chains in addition incorporate halo effect around individual atoms to improve the depth-cueing. The animation allows to zoom in and out, as well as other tools such as VMD [12].

In the case of the tyrosine protein, for example, we are interested in an important residue (Phe121). An inhibitor is a molecule that binds to a protein and prevents it from functioning. Inhibition means that they have complementary structure (also referred to as a lock and key model). The investigation of the detailed structure of such a complex, when a disease is caused by a malfunctioning protein, for often important to find compounds which would have the necessary characteristics to be inhibitors of the malfunctioning protein. To do this, an in-depth 3D analysis of the structural aspects is very important.

The new two-level rendering approach enables us to investigate the protein in their full complexity (with all atoms), even if they are composed of hundreds of thousands of atoms, interactively in 3D (see Fig. 6a). We are able to investigate their detailed shapes while they are moving smoothly. When interacting with the protein, we regularly use a temporary switching to 2D view, for example, to see the structure of the molecule in the current view - due to its complexity and large number of atoms it is not always immediately obvious which parts connect to which others along the backbone. With our new rendering approach we can very easily suppress the rendering of residues and thereby only show the backbone of the protein (see Fig. 6b). In some situations, however, for a very detailed view, we temporarily disable the rendering of all atoms but just those of an selected residues (Fig. 6c-f).

Another protein we are interested in is the protein tyrosinase (TYR) [14]. It has a relatively complex structure and showing it with all atoms results in a pretty packed visualization. For an overall analysis, we again make use of the option to not show residues but only the backbone. To better explore the structure of the backbone we apply an interactive 3D fish-eye distortion (with care, however, i.e., moving the fish-eye around slowly). The fish-eye view locally separates parts of the backbone and thereby reveals details (see Fig. 8).

Five experiments indicate that interactive visualization in our VR lab protein structures is conveyed much clearer from stereo projection as opposed to desktop environment. Figure 7 demonstrates interactive visualization of tyrosine protein with its inhibitor BPTI (Bovine Pancreatic Trypsin Inhibitor). Additional material is available on <http://www.cmc.uio.no/research/protein-dynamics/>.

### 5 PERFORMANCE ANALYSIS

We have tested the performance of our two-level rendering approach on the chequerboard GroEL (LAC9) consisting of 56,084 atoms for NMA and several other (see Table 2) with 1000 force calculations.

While rendering the GroEL NMA analysis we achieve 29 FPS with 3200 x 1200 resolution in the stereo mode (or 58 fields per sec). For the results are presented in Table 3 and Figure 9. These results were made using a NVIDIA GeForce GTX graphics card. The stereo results were made on the NVIDIA Quadro graphics card. The stereo results were made by implementing linear calculations on the GPU (once per atom) and a significant factor in our gain in performance lies in that with geometry offset to once per atom vs four times per atom without the geometry shader.

We compare the bandwidth load between our approach and one-level approach also using halfwidths for object representation but not for geometry generation. Table 2 shows that the average atom count per residue lies between 7 - 9 for proteins when we use the following formula for the bandwidth in bytes per sec:

One-level rendering:

$$\left( \frac{\text{atoms} \cdot \text{vertex} \cdot \text{float} \cdot \text{byte}}{\text{atoms} \cdot \text{vertex} \cdot \text{float}} \right) = (\text{atoms} \cdot 4 \cdot 3 \cdot 4) \quad (9)$$

Two-level rendering:

$$\left( \frac{\text{residues} \cdot \text{floats} \cdot \text{byte}}{\text{residues} \cdot \text{floats}} \right) = (\text{residues} \cdot 6 \cdot 4) \quad (10)$$

It turns out that our method uses approx. 15 times less bandwidth as compared to a one-level approach. Moreover, this comparison does not include rendering of hydrogen atoms as they are usually neglected in molecular visualization. In cases where protein visualization would require rendering of hydrogen atoms, the atom count will increase, while the residue count will stay the same. This means that the bandwidth and calculation load for one-level approach will be at least doubled, whereas our load will stay the same.

### 6 SUMMARY AND CONCLUSIONS

We have designed a two-level rendering approach for interactive visualization and exploration of protein dynamics. Our approach performs significantly less bandwidth as compared to a one-level approach, that renders all protein. During studies of protein, biologists often like to switch between different levels of abstraction in protein, e.g., first we might be interested in getting the big picture and seeing the backbone structure only.



Fig. 7. Photograph taken from a visualization of structure of tyrosine protein with its inhibitor BPTI (Bovine Pancreatic Trypsin Inhibitor) in our interactive environment. Photo: © Espen Erik Larsen, www.bcf.uio.no

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8

# Summary, Conclusions, Outlook, Acks

- Few sentences on what has been achieved
- Conclusions is very important section!
  - What have you learned during the research
  - What should be people aware of
  - What turned-out to be most difficult
  - Under which assumptions the approach is valid
- Usually written in the past form
- Outlook
  - Further potential application areas
  - Improvements
  - Plans
- Acks: thx for data, illustrations, brainstorming



# Thanks goes to

- Meister
- Helwig Hauser
- Thomas Theussl
- Werner Purgathofer
- Robert Tobler
- Dieter Schmalstieg
- Christian Breiteneder
- George Gopen and Judith Swan

# Further Reading

- How to write a scientific paper  
<http://www.cescg.org/guidelines/>
- Scientific Writing (in German)  
[http://www.ims.tuwien.ac.at/teaching/gma/nachlese/wiss\\_schreiben.php](http://www.ims.tuwien.ac.at/teaching/gma/nachlese/wiss_schreiben.php)
- The Science of Scientific Writing  
<http://www.amstat.org/publications/jcgs/sci.pdf>
- Great page about giving a talk and writing a paper  
<http://www.wit.at/events/peyton-jones/>