Structures of Life: The Role of Molecular Structures in Scientists' Work

Dhaval Vyas[1], **Olga Kulyk**[1], **Paul van der Vet**[1], **Anton Nijholt**[1] and **Gerrit C. van der Veer**[1, 2]

Human Media Interaction Group[1]
University of Twente
Drienerlolaan 5, 7522 NB
Enschede, the Netherlands
{d.m.vyas | o.kulyk | p.e.vandervet | a.nijholt}
@ewi.utwente.nl

School of Computer Science[2]
Open University
Valkenburgerweg 177, 6419 AT
Heerlen, the Netherlands
gerrit@acm.org

ABSTRACT
The visual and multidimensional representations like images and graphical structures related to biology provide great insights into understanding the complexities of different organisms. Especially, life scientists use different representations of molecular structures to answer biological questions and to better understand cellular processes. Combining results from two field studies, we explore the role of molecular structures in life scientists' current work from a human-factors perspective. Our main conclusion is that different representations of molecular structures, due to their visual nature, are important for supporting collaboration, constructing new knowledge and supporting scientists' professional activities in general.

Keywords
Life sciences, molecular structures, HCI, design.

INTRODUCTION

> *"To understand biology, one must think in a language of three dimensions, a language of shape and form… in biology, especially at the cellular and molecular levels, nearly all activity depends ultimately upon form, upon physical structure..."*
>
> John M. Barry (2004, p.100-101)

Life is currently understood as being constituted by processes at the molecular level. The molecules and molecular assemblies involved in making cells work are incredibly complex. The structure of molecules is an important key to their properties. Structure can only be understood in three dimensions. Moreover, the structure of a molecule is vastly different from the structure of more familiar objects like tables and chairs. A molecule's structure must be understood in relation to its surroundings and the interactions with the surroundings are dynamic. To make matters even more complicated, many molecules of biological importance can undergo changes in form, which in turn changes their properties. Understanding these processes is a key to understanding the functioning of living cells, both in the normal or healthy state and in the diseased state. It can help to understand why current drugs work and how they work and to identify drug targets. The study of biological molecules is therefore of great economic importance as well. For us, the important lesson is that visualizations of molecules play a key role in the understanding by scientists of molecular function.

There are a very few published studies of human-factors support in life science research, for example, user evaluation of existing visualization tools (Saraiya et al. 2004), task analysis for studying work patterns of bioinformatitians (Bartlett and Toms, 2005), measuring users' performances while interacting with molecular structures in virtual environments (van Liere et al. 2005). In this paper we focus on understanding the role of molecular structures in the practice of professionals working in different fields of life sciences. By the term 'molecular structures' we mean different possible visual representations of the structures of molecules of biological importance, such as proteins and RNAs. We intend to understand the use of different representations of molecular structures and how these representations help life scientists in their everyday work.

We combine the results of two field studies related to understanding life scientists' work practices. We carried out a set of contextual interviews, observations, diary study and questionnaires on professionals from different fields of life sciences. In this paper, we specifically focus on the molecular structures and show that representations of molecular structures, due to their visual nature (e.g. form, color), are an important means for supporting collaboration, constructing new knowledge and carrying out other professional activities.

In the following sections, we first provide a short background on different visual representations of molecules, to establish a clear understanding of their importance. Next, we describe our field studies and discuss the results.

THE USE OF MOLECULAR VISUALIZATIONS
Chemists and biologists use visualizations of molecular structure for a variety of tasks that are closely related to the research in which they are involved. The importance of molecular visualizations in these disciplines can hardly be overrated. A large variety of visualizations

has emerged and many software packages allow one to easily switch between different kinds of visualization. Figure 1 shows three often used visualisations.

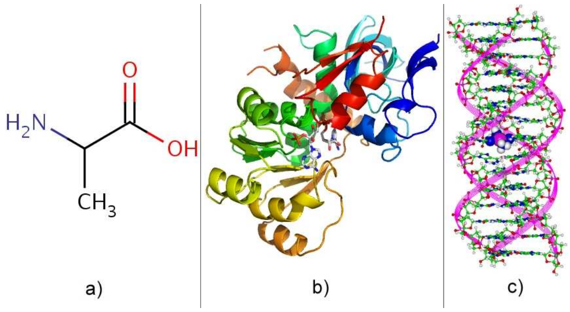

Figure 1. Different molecular structures: a) amino acid alanine, b) protein p53 tumor suppressor, c) part of DNA.

Figure 1(a) shows a planar graph of a possibly three-dimensional molecule (certain atoms are left out for clarity). This kind of visualization, called a molecular structure, helps chemists to quickly identify what kind of molecule is involved and what possibilities for chemical interaction it has. This kind of visualisation is only useful for smaller molecules. A similar visualization of a protein, which typically consists of thousands of atoms, would be hard to interpret.

Figure 1(b) shows the secondary and tertiary structure of a protein in the so-called cartoon view. A cartoon view directs the viewer's attention to important structural elements like α-helices, β-sheets, loops, and turns. A cartoon view comes close to a representation of the 3D shape of a protein while at the same time structural elements are visible. Cartoon views are used by molecular biologists to explore the properties of proteins, in particular where they can bind and to which species they can bind. This information, in turn, can help to learn more about the function of the protein. Because the representation is 3D, the details at the back are obscured. Therefore, a good visualization tool allows its users to rotate the image around its three axes. In fact, it is observed (Kulyk and Wassink, 2006) that users always start to rotate these views when they explore a new protein. A cartoon view is not enough, however; other views, such as the van der Waals surface view and the solvent-accessible surface view, are needed as well. In addition, users want to zoom in and out; and the tool should display chemical information about selected regions, such as which amino acid residues are present in the selected region. Modern protein visualization tools enable their users to easily switch between views.

Figure 1(c), finally, shows a visualization of a part of a double DNA strand. This kind of visualization is commonly known as ball-and-stick. Atoms are represented as balls, chemical bonds as sticks between balls. The aim is didactical at approximately the level of secondary school pupils. The helical nature of the double strand is highlighted and the fact that the two strands are held together by hydrogen bonds (not visible as such in this figure) is clarified. Atom kinds are indicated by colors. The idea is not to visualise DNA as it is present in the cell nucleus.

FIELD STUDIES

We carried out two explorative field studies (Vyas et al. 2006; Vyas et al. 2007; Kulyk et al. 2006) of life scientists from various disciplines (biology, chemistry, and bioinformatics) and with various levels of expertise. Our studies included contextual interviews, diary studies, ethnographic observations and questionnaires, over a period of 3 months. The first study included 30 scientists and the second had 57. The overall purpose of our studies was to explore working practices and experiences of scientists in their real-life settings.

During the contextual interviews, the questions were focused on understanding scientists' research activities, their workflow structures, use of different tools, collaborative activities, and domain-related problems. These interviews were carried out mainly on-site but in certain cases also remotely using the WebEx tool. In the diary study, the scientists were asked to log in the details of their daily habits and main activities for at least 10 days. In the diary a set of questions were also provided in order to collect specific information. We also followed the scientists and observed their laboratory sessions where our team spent a full day with the scientists to observe their work flow and day-to-day interactions. Our questionnaire was based around knowing how the scientists deal with their domain related problem.

THE ROLE OF MOLECULAR STRUCTURES IN LIFE SCIENTISTS' WORK PRACTICES

We discuss our findings about the molecular structure representations in the following.

Collaboration

Figure 2. Collaborative work in the laboratory.

Collaboration is an essential aspect of life scientists' work, ranging from working together in the same physical space to co-publishing and distantly sharing work. During our field study we observed that being aware of each others' activities was an important aspect of scientists' work in the laboratory (Kulyk et al. 2007). Bioinformatics laboratories are full of displays (Figure 2) that show different visualizations and other relevant visual information. Scientists use these representations for analysis, discussion, comparison, decision-making and for refining their experimental strategies.

Because of the multidisciplinary nature of life science research, effective communication between team members is an important part of the project's overall success. These research teams normally involve researchers with backgrounds in biology, bioinformatics, chemistry, mathematics and statistics. In this kind of collaborative work there are work dependencies, i.e. scientists and other professionals have to base and adapt their work to suit others. Scientists in this kind of teams communicate with each other by using different forms of images, visual structures and other types of visual information like graphs and charts. Because of their visual nature, molecular structures played a part for providing a platform for communication. As one of the scientists commented while discussing his work, "*these images become a crucial communication tool. These graphical images help us explain to them* [other colleagues from different discipline] *what we exactly want. Because of these computationally generated graphical structures I can show my colleague by pointing to a specific portion of a cell."*

Constructing new knowledge

Our field study showed that scientists construct new knowledge based on their collaborative expertise and goals. They always generate new questions and queries to solve biological problems. Molecular structures play a role in gaining new insights and generating new cellular models. Molecular structures are used to simulate certain situations and to predict the behaviour of molecular interactions that evolve in cells, for example, in simulating protein interactions.

We frequently observed during our field studies that being able to change, manipulate and handle different visualization styles of molecular representation was a key to understanding functioning of the cell. This continuous interaction with visual representations, including the ability to modify them, helped scientists to develop better understanding of the cell structures. Thus, over time scientists could become capable to efficiently make sense of cell structures.

Supporting professional activities

We observed that molecular structures were heavily used during experiments and simulations. One of the scientists commented during our contextual interview that "*once you know how these proteins and DNA are* [physically] *arranged, you have a much better sense of their function*." We observed that molecular structures play a major role in the analysis and interpretation of life science experiments (Kulyk and Wassink, 2006). In several cases, we observed that the reason for creating and manipulating the representations of molecular structure was to provide evidence for and results of life scientists' hypotheses. As some scientists expressed, graphical images and representations are the outcomes of their research. Importantly, scientists use various types of representations to support their ongoing research. As one of the scientists suggested, *"all of these structures bear different levels of importance and bear different kinds of information. These all depend on the aim of the research and method that we want to apply in our research."* Especially researchers with a biology or chemistry background need to look for chemical structures and chemical properties. Interestingly, the role of these representations was also seen outside the laboratories. Scientists working in academia use these representations in writing their papers, teaching, and presentation activities. These scientists use images, graphs, tables and other data related to the molecular structures for presentations and teaching.

LESSONS LEARNED

The results of our studies confirm Kidd's (1994) argument that computer support for knowledge work should be targeted on the '*act of informing*' rather than passively providing information in a disembodied form. Molecular structure representations are about connections, specific shapes, and the forces that govern the form and the interaction with other molecules. We believe that informing scientists and researchers about different ways and modes to interact with the visual representations should be given prime importance. Based on our analysis we indicate that there are some lessons to be learnt from this exploration.

Providing several viewpoints for analysis

Drawing conclusions and making decisions based on a certain structure is of paramount importance for life scientists. Especially when a multidisciplinary team works in a collaborative fashion on a subset of visual information, several different ways of representing structures are required to enhance exploration of heterogeneous information and to support cooperation between disciplines. We believe that these scientists should be able to generate specific viewpoints of molecular representations based on their specific needs. Creating, exploring and changing molecular structures in novel ways would allow scientists to inform each other of new insights.

Scientists should be able to manipulate, transform, scale, and move a specific part of a structure in intuitive ways. Designers should provide different analytical viewpoints and links between these viewpoints, which could allow scientists to make better judgments about structures. Designers can think of new ways of visualizing and representing information, providing support in such a way that scientists can keep track of their different ongoing processes, allowing annotations on different visualizations that can be used next time when they work together in a team.

Making structures more explicit

We observed that scientists tend to make quick judgments about the properties of structural elements of a molecule and what they mean for the properties of the whole molecule. When these structural elements are made explicit, e.g., by highlighting a specific part of the representation of a molecule, it helps scientists in searching for specific insights. By making specific parts of the representations explicit, annotations can be facilitated in different ways (e.g. using touch-screens), and even co-annotations amongst collaborators are possible. These annotations are not only useful to understand certain facts but they also serve as cue to go back to previous phases of reasoning and discussion (Kulyk et al. 2007).

DISCUSSION

This work should not be seen as a full-fledged, complete account on the usefulness of molecular structures in the work of life scientists. In fact, it can be seen as a way of making sense of advanced biological research. One of the main claims of this paper is that the role of molecular structure representations in life scientists' work is three-fold: support of communication and collaboration in a multidisciplinary team, help in creating new knowledge about molecules, and support of their professional activities.

We were especially interested in getting a larger picture of molecular structures by understanding the 'enactment' of these scientists in their working environments. What was observed throughout our exploration was that the visual information and different visualization practices of life scientists played an important role. Overall, molecular structures and other visual information supported by visualization techniques and computer-generated images help scientists in the following aspects:

- They provide evidence and results of their research. As several participants expressed, these images and structures (also) serve as the outcomes of their research.
- Moreover, these 3D (or 2D) visuals help to better understand what is happening at the micro- and molecular level.
- They help in communicating and presenting the results in a desired way so that other people can make sense of it.
- They are also a collaboration tool for the colleagues and collaborators to understand what a scientist wants from them. This eventually improves the working process.

CONCLUSION

The main reason to bring the research related to molecular representations to the HCI community is that this field of research has not been well studied from a human-factors point of view. Especially since research programs on molecular biology and bioinformatics promise to yield cutting-edge solutions to life-threatening problems (e.g. HIV, cancer) in a less expensive way, we believe that a human-factors perspective will only improve the current understandings of biological aspects within a larger research scope.

Representations of molecular structures are vital to life scientists' work practice. Our field studies show that, due to their visual nature, the representation of molecular structures supports collaboration, construction of new knowledge for didactical purposes, and in general, supports scientists' professional activities.

REFERENCES

Barry, J. M. (2004) The Great Influenza the epic story of the deadliest plague in history. New York: Viking Penguin.

Bartlett, J.C. and E.G. Toms. (2005). Developing a Protocol for Bioinformatics Analysis: An Integrated Information Behavior and Task Analysis Approach. Journal of the American Society for Information Science and Technology, 56(5): 2005, 469-482.

Kidd, A. (1994) Marks are on the Knowledge Worker. Proc. of CHI'94. ACM Press: NY, 186-191.

Kulyk, O. and Wassink, I. (2006). Getting to know Bioinformaticians: Results of an exploratory user study, in British HCI Workshop Combining Visualizations and Interaction to Facilitate Scientific Exploration and Discovery, September 12 2006, T. Adriaansen and E.V. Zudilova-Seinstra (eds), London, UK, 30-36.

Kulyk, O., van Dijk, B., van der Vet, P.E., & Nijholt, A. (2007) Do you know what I know? Situational Awareness and Scientific Teamwork in Collaborative Environments. In Proc. of SID'07, CTIT, Vol. WP07-02, Enschede, 207-215.

Saraiya, P., C. North, and K. Duca. (2004) An Evaluation of Microarray Visualization Tools for Biological Insight. in INFOVIS '04: Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'04). Washington, DC, USA: IEEE Computer Society.

van Liere, R., Kok, A., Martens, J.-B. and van Tienen, M. (2005). Interacting with Molecular Structures: User Performance versus System Complexity. In Eurographics Symposium on Virtual Environments, (Aalborg University, Denmark), 147-156.

Vyas, D., de Groot, S. and van der Veer, G.C. (2007). Searching and Archiving: Exploring Online Search Behaviors of Researchers. Proc. of HCII'07, LNCS, Springer- Heidelberg, 360-364.

Vyas, D., de Groot, S. and van der Veer, G.C. (2006). Understanding the Academic Environments: Developing Personas from Field-Studies. Proc. of ECCE'13, ACM Press: NY, 119-120.