Improving Student Reasoning on Molecular Interactions Using the Context of Drug Development in Chemistry Education

Laura van den Bosch | 3961125

Paper research project | 3 July 2019

Utrecht University

Research Project SEC:	30 ECTS
Supervisor:	dr. G. T. Prins (assistant professor) Freudenthal Institute for Science and Mathematics Education, Faculty of Science, Utrecht University, Utrecht, The Netherlands.
Journal type:	Chemical Education Research and Practice (CERP)
Guidelines:	The goal of this journal is to bridge the gap between the two groups so that researchers will have their results seen by those who could benefit from using them, and practitioners will gain from encountering the ideas and results of those who have made a particular study of the learning process.
Audience:	Chemistry Education researchers and teachers

Improving Student Reasoning on Molecular Interactions Using the Context of Drug Development in Chemistry Education

L. van den Bosch, July 2019

Abstract

The goal of this study was to evaluate to what extent a curriculum unit helps students to improve their reasoning concerning molecular interactions. Chemical models are very important in order for students to make sense of the abstract concepts of molecules and their interactions. An authentic practice, in this study a drug development case, is expected to enhance student understanding and reasoning. Furthermore, ICT-tools can be helpful in the classroom to help students better understand molecular interactions. A curriculum unit for the upper grades in high school (5vwo) was designed that focuses on molecular modeling with ArgusLab. Students were challenged to find an alternative drug molecule for malaria, applying their chemistry knowledge concerning molecular interactions such as hydrogen bonds. By creating worksheets and a think aloud protocol, data was obtained and analyzed using a measuring instrument that scored their reasoning level from 1 to 4. This measuring instrument was created based on expert interviews. It can be concluded that students improve in their reasoning from level 0 or 1 towards level 2 or 3. The curriculum unit helps students to improve their understanding of molecular models and their interactions, which confirms previous results on authentic practices in molecular modeling. Further research should focus on using this knowledge to develop modeling practices in other chemistry subjects and specifically studying the effect of the use of an ICT-tool such as ArgusLab.

Introduction and theoretical background

Models in science

Models in general. Science tries to provide explanations for phenomena in the natural world. For that purpose, descriptions and simplifications of complex phenomena are made in order to make sense of what is happening on a macro level. These descriptions and simplifications are called 'models' (Gilbert, 2005). In order for students to gain a basic knowledge of science, a central skill is to be able to make sense of those scientific models. Models are necessary to grasp scientific concepts (Giere, 1988), but are also very hard for students to give meaning to (Grosslight, Unger, Jay & Smith, 1991). Every student has created his own mental models about phenomena he or she observes in his or her surroundings. Mental models can be explained as a mental representation which people automatically generate (Vosniadou, 1994).

Models in chemistry. This study will focus on the use of models in chemistry education. Chemistry is generally considered a hard and abstract discipline and many studies have been performed to improve chemistry education (see e.g. Roberts & Bybee, 2014; Holbrook & Rannikmae; Hurd, 1998). Harrison & Treagust (2005) researched the mental models secondary education students hold concerning atoms and molecules and the problems that arise in those models. They also list seven categories of models that are used in science or specifically in chemistry. The analogical model is the most important model that we will refer to here: a ball-and-stick molecular model, for example, can be considered an analogical model, because it represents the properties of a molecule without being built in the correct proportions. The study reports on conducted interviews with students and categorizes them in levels of modeling ability, based on the levels Grosslight et al. (1991) defined, where level 1 modelers tend to interpret the

models as partly literal: they mistake the model for reality. Level 3 modelers are more advanced in their understanding and reasoning about models. Harrison & Treagust (2005) conclude that most of their interviewed students are level 1 modelers. Their recommendation is to build modeling skills with students, defining those skills as "understanding of spatial molecular structures and the ability to transfer between molecular representations and chemistry understanding levels." (Dori and Kaberman, 2012). Improving these skills will enhance their understanding. It is especially important to show them where analogical models break down and what the shared characteristics of the model with the target are. Students need to be able to translate drawn 2D models to 3D models; this is an important step towards a better understanding and interpretation of chemical concepts (Dori and Kaberman, 2001).

Molecular modeling

Authentic practices. These molecular modeling skills can be enhanced by using authentic practices: Using an authentic real-world modeling context is expected to promote a deeper understanding of models and their use in science education (Golbert & Pallant, 2004). When the learning of models and modeling can be legitimized for a student, he or she will recognize the relevance of learning chemistry and its models (Prins, Van Driel & Bulte, 2008).

Drug development. A specific application of molecular modeling in research is drug development. A commonly used technique in the process of drug design is structure-based virtual screening (SBVS). This technique provides in the need for a fast search for small molecules that may bind to specific targets. A 'docking program' can be used to place a small molecule into, for example, the active site of an enzyme. The energetically most favorable conformation and orientation can be found by 'scoring' the molecule. One can proceed by performing a lead optimization: the process where an existing molecule is modified in order to further lower the

binding energy of the interaction with the binding site of the enzyme, thus blocking its activity. This technique can be used when there is already a lot of information available on both the binding site and the drug molecule (Jorgensen, 2008).These are generally the first steps in a process which can then be followed by specific selection of molecules for further (in vitro) research (Kroemer, 2007). We chose the drug development case as an authentic practice for this research.

Malaria disease. The strategies mentioned above are used very generally in drug development research. In this study, we chose to zoom in on malaria disease. Malaria afflicts 500 million people worldwide and is therefore a big problem. However, the parasite enzymes are becoming resistant to the existing antimalarial drugs. These drugs specifically act by inhibition of an important parasite enzyme (DHFR), that plays an important role in the replication. One of the drug molecules that has been researched extensively is pyrimethamine. Due to mutations, the enzyme cannot be inhibited by pyrimethamine anymore (Peterson, 1988). Using lead optimization, an analogue may be available that does block the mutated enzyme.

ICT-tools for visualization

In drug development research, ICT-tools are used for the purpose of visualizing molecular models and calculating their interactions. Dori and Barak (2001) studied the use of physical and virtual (computerized) molecular models in the classroom in order to improve students' understanding of new concepts and concluded that these models and visualizations in general are very important for reaching this goal. They state that both ways of representing molecular models have advantages; the advantages of the computerized models are that the possibilities in terms of various colors, model types (ball-and-stick, space-filling) and sizes are unlimited, but also that virtual models can be used to perform calculations in order to minimize energy and predict 3D-shapes (Dori and Barak, 2001). Computerized modeling; the visualization

of (complex) molecules and their interactions, can furthermore help affecting the mental structure of models students have. Viewing 3D animations can improve students' incomplete mental models of the dynamic nature of chemical reactions (Sanger et al., 2001). Dori and Kaberman (2012) state that the difficulty for students is to connect the macroscopic properties of matter with their microscopic explanations. The researchers conclude that using the virtual learning environment improves the modeling skills of students at all academic levels. Students gain a better feeling for the corresponding between molecular formula, structural formula and 3D structures.

Current study

Curriculum unit. In the case of this study, a curriculum unit will be used, where students get in the role of a researcher who is developing a drug against malaria disease. Previous studies show this is an interesting context where students can apply their knowledge of chemical bonds and interactions on an authentic and challenging problem (Bode, 2016). An ICT-tool called ArgusLab will be used to let students perform a lead optimization of pyrimethamine. There has been research on 3D visualizations and how ICT-tools help students to understand models better, but to our knowledge, no studies have been done on student's reasoning on molecular interactions, using a context-based molecular modeling approach.

Research question. In this paper, the results will be presented and some implications for this molecular modeling practice for chemistry education will be explained. The research question that this study aims to answer is:

> To what extent are students able to achieve a higher level of reasoning on molecular bonding/interactions during this context-based curriculum unit with a molecular modeling approach?

Methods

In this section, the revision of the curriculum unit 'Molecular Modeling, Drug Development against Malaria Disease' is explained. The research setting along with the lesson plan and a data collection plan will follow, which is in turn followed by a description of the measuring instrument created to analyze the data.

Curriculum unit

The unit (summary in Appendix A) that will be used in this study has been developed to fit the curriculum of the upper grades of chemistry, where it is specified that students should be able to explain the contribution of molecular modeling in science research. Given that the existing Dutch chemistry textbooks do not pay attention to this topic, the development of this unit was necessary.

The unit has been tested among grade 11 (5vwo) students who are taking chemistry. The results have been promising, but improvement is necessary, and some recommendations are done in order to do that (Bode, 2016). Based on studies mentioned before, we expect the unit to improve the student's reasoning on molecular interactions, due to the use of an authentic practice (Golbert & Pallant, 2004), and a molecular modeling program (Dori and Barak, 2012), called ArgusLab. ArgusLab is used to visualize proteins and ligands and to calculate the binding energies of different pyrimethamine analogs. Various activities are included to let students do research towards drug design and malaria. Experts were consulted in order to make sure that the unit actually represents a real-world context and is conceptually correct, and to check that the strategies that are used by experts correspond with those we expect students to learn.

The main finding from both interviews is that students should work with a specific plan when designing a drug. They should choose one step and use their knowledge to reason why they

choose to change that specific characteristic group – every step should be deliberate, trial and error is not acceptable. For example, a fluorine atom can be exchanged for a chlorine atom and one could compare those two. Students should furthermore never vary two variables at the same time. Using bioisosterism is an important part of the strategy that experts use. We can ask students to exchange an NH-group for a dimethyl group and then ask them what is being lost by doing so: hydrogen bonds, while you get hydrophobic interactions back. Also, students can reflect upon what parts of the molecule should not be changed and why.

To collect data, specific questions were added and adapted. Worksheets were created in order to be able to analyze the answers students wrote down.

Research setting

Lesson plan. The curriculum unit was carried out at the GSR in Rotterdam in 5vwo in three weeks, between April 1 and April 18, 2019. There were two blocks of 100 minutes per week in the first two weeks, and one block in the third week. For a more detailed lesson scheme, see Appendix B. Since this was a small scale exploratory pilot, no HLT was created.

Data collection

In the curriculum unit, three specific moments are chosen to use as a data collection component. In this section, these three moments are specified and their goals explained (see also Table 1).

Measuring moment 1. The first moment is also called the pretest (assignment **6d**). It is placed after the introduction of the unit and the chapters on malaria and right before the students actually start the docking assignments. Prior knowledge concerning amino acids, enzymes and

proteins has already been activated at that point, so the pretest measures to what extent they can already apply their knowledge in this context.

Measuring moment 2. During the second moment, nearly at the end of the unit, the students are asked to design their own pyrimethamine analog and reason about how they do that and why they change various characteristic groups, for example using the bioisosterism strategy. They have to draw their design and also write down their argumentation.

Think aloud protocol. During the third measuring moment, a think aloud protocol will be applied. The students are recorded and asked to explain their thoughts while they think about a new pyrimethamine analog (Appendix D)

Table 1

Data components

Title	Description	Goal
Measuring moment 1	A question in the unit (**6d**), asking students to write down what they would change about pyrimethamine in order to make a useful drug molecule.	Pretest: check what level of reasoning students use before they are taught about bioisosterism and docking.
Measuring moment 2	A question in the unit (**final assignment**), asking students to design a pyrimethamine alternative and dock it in ArgusLab.	Part of the posttest: let students write down their reasoning and draw interactions.
Think aloud protocol	Their thoughts when they are performing the final assignment are recorded.	Part of the posttest: let students explain their thoughts and argumentations.

Data analysis

In this section, the data analysis method is explained. One of the conclusions of the previous study was that it was hard for students to visualize and understand the 3D model of the

enzyme with the inhibitor; and it was hard to measure if their understanding improved after doing the assignments. For the purpose of analyzing the data gained from students, a measuring instrument was created, used to 'score' the student's answers and thoughts. Expert interviews were conducted to get an idea of their reasoning.

Measuring instrument. For the purpose of creating a realistic measuring instrument, experts were interviewed to create a 'top level'. What stands out in the interviews is that experts reason about a lot of different interactions when they talk about designing a new drug. It turns out that experts are able to name hydrogen bonds, but also take into account ion-ion interactions (or explain why those might not be present) and Vanderwaalsforces. They are able to use pi-pi stacking of two benzene rings (this is not part of the high school curriculum, so it is not likely that students will mention this). It is very clear from the interviews that experts use a worked out step-by-step strategy for the purpose of modifying a molecule and that they specifically exchange characteristic groups in order to create or delete an interaction at that location. A measuring instrument matrix has been designed with four levels, with level 4 being the 'expert level', based on the strategies and interactions mentioned by the experts.

In level 4, it is expected to name at least three important interactions and reason about their relative importance in the binding of the ligand to the protein. Also, students need to be able to connect these interactions to characteristic groups of the analog, they have to draw and localize the interactions between the ligand and the binding site and they need to have a clear plan and reasoning about what they are changing in the molecule and why. They are expected to use the bioisosterism strategy in a correct way. Level 1 is then defined as a level where students at most name one interaction, without connecting that to a certain characteristic group. It is expected that students functioning at level 1 use trial and error instead of a step-by-step strategy.

In the horizontal part of the matrix, six interactions are distinguished that could be named by students, based on the interactions experts use. Every answer in the three measuring components is scored using this instrument. This results in a level indication in the pretest and posttest components, which can then be compared. The measuring instrument can be found below:

Table 2 Measuring instrument

	Sterical hindrance/ hydrophobic interactions.	Pi-pi stacking	Hydrogen bonds	Ion-ion interactions	Vanderwaals forces	Dipole-dipole
Level 1 1. Students are able to name one interaction. 2. Students are not able to connect these interactions to specific characteristic groups of (analogs of) pyrimethamine. 3. Students are not (or barely) able to localize the interactions, that is, the interactions between characteristic groups of (analogs of) pyrimethamine and the amino acids of DHFR's bindingsite. 4. Students seem to use trial-and-error when finding a new analog. 5. Students mainly use wrong terminology like 'matter contains hydrogen bonds' instead of 'can form hydrogen bonds'.						
Level 2 1. Students are able to name **at least** two interactions. 2. Students are able to connect these interactions to specific characteristic groups of (analogs of) pyrimethamine. 3. Identical to Level 1 point 2 4. Students seem to work according to a plan to some extent. 5. Students use correct terminology.						
Level 3 1. Students are able to name **at least** three interactions, and they seem able to reason about the interactions (e.g. strongest) 2. Identical to Level 2 point 2 3. Students are able to localize the interactions, that is, the interactions between characteristic groups of (analogs of) pyrimethamine and the amino acids of DHFR's bindingsite and can thus say which orientation of the molecule will fit the best. 4. Students work according to a plan and change characteristic groups purposefully. 5. Identical to Level 2 point 5						
Level 4 1. Students are able to name **at least** three interactions, and they are able to reason about the relative strength of the interactions 2. Identical to Level 3 point 2 3. Identical to Level 3 point 3 4. Students work according to a plan and change characteristic groups purposefully. They use the bioisosterism strategy correctly. 5. Identical to Level 3 point 5						

In order to apply the measuring instrument on the student data, all answer sheets are divided into the four groups (A-D) and for every group, the think aloud protocol recordings are transcribed. In those answersheets and the transcripts, the answers are scored, answering the following questions:

1. Which interactions are mentioned?

2. To what extent are students able to reason about the interactions they mention?

For every group, the measuring instrument is filled out, resulting in a level of reasoning. In the results section, the results of using this instrument will be presented.

Results

The observations during the implementation of the curriculum unit indicated that the educational process was carried out with sufficient quality. The teacher was satisfied with the content and it was generally clear for students what they had to do. Four groups of two students completed all three measuring components. We believe this setting was appropriate to use for this research. In the conclusion and discussion section, improvements and recommendations for future implementation will be given. During the pilot, several assignments were skipped due to a lack of time. However, these assignments were considered of less importance for the goal of the unit, so we believe that this did not influence the results.

Measuring moments

In this section, the results for the components will be presented in two parts:

1. A pretest component (*measuring moment 1*)
2. A posttest component (*measuring moment 2* and *think aloud*). The latter will consist of written results of the last assignment and the results of the verbal think aloud protocol.

For both the pretest and the posttest, the results will be presented in a table, followed by an explanation based on quotes from the written and verbal data.

Pretest

Table 3

Pretest results from written answers

Group	Level	Interactions mentioned
A	0	-
B	1	Covalent bonds
C	1	Hydrogen bonds, covalent bonds
D	1	Hydrogen bonds

This pretest question was answered very generally by the students. Two of the groups mentioned hydrogen bonds, but did not mention where in the molecule those hydrogen bonds could be formed and by what kind of group.

One of the groups referred to disulfide bridges as a possible new connection between the molecule and the binding site.

> Group C: "You have to make sure that there will be more attraction, by letting it form disulfide bridges or hydrogen bonds."

Group A felt like they did not have the knowledge necessary to answer this question.

> Group A: "I can not comment on this usefully due to a lack of knowledge."

Posttest

Table 4

Posttest results from written answers and think aloud protocol

Group	Level	Interactions mentioned
A	1	Hydrogen bonds
B	3	Hydrogen bonds, hydrophobic interactions, electronegativity (dipole-dipole)
C	2	Hydrophobic interactions (sterical hindrance), dipole-dipole
D	2	Hydrogen bonds, vanderwaals forces, hydrophobic interactions

For these results, the two components of the posttest have been analyzed and scored. For a group to be scored in a certain level, the answers of both the components needed to fulfil the requirements of that level. The group that did not have one idea during the pretest, improved to level one. The group did still find it very hard to reason about a design and tended to strand on side tracks instead of focusing on the assignments themselves.

> Group A: "What do these R-groups mean? You need to change things in those R-groups in order to make bonds, right?"

> Group A: “Let’s just add an OH-group or something! Check if it makes a hydrogen bond... No idea how you can see that, but... [...] I propose we just add something at Y that can form a hydrogen bond. Check if it forms a hydrogen bond with whatever group in there.”

Group B did actually improve a lot. This was a group that seemed enthusiastic from the beginning. In the first two lessons, they finished their assignments very fast so they could already start working with ArgusLab. In the second moment and the think aloud recording, they were scored in level 3. They started out using more or less trial and error and reading from the curriculum unit on bioisosterism:

> Group B: “If we change all this to F-atoms.. Well, in the manual it says that it has electron sucking characteristics. And then we can change this one back from S to Cl.”,

When asked, however, they could reflect on where they thought a $(CH_2)_5CH_3$ chain would fit in the binding site, what happens if it gets too big and they also tried to draw that on the worksheet.

> Group B (upon asking what the green line indicating hydrophobic surroundings in PoseView images means): “So we could place a hydrophobic chain there... [...] You can add extra C’s, to make sure it makes a hydrophobic chain. So you can exchange some of those H-atoms for C-atoms.”

Group B also tried to reason about the relative influence of various changes:

> Group B: “You could make a triple bond to an N-atom. That has the same effect, but I think that it will not have the same mass change. And also, you will have a hydrophobic character, so that might help.

Group C had a hard time finding out how exactly they should use the program and what the exact assignment was. They really focused on dipole dipole bonds, using fluorine atoms. They were scored in level 2, because they were able to reflect on the interactions in the analog molecule, but failed to make the connection between binding site atoms.

> Group C: "Because that group [the CH_3-group they added] is hydrophobic and does not want to bind, so it will kind of turn and then the fluorine atom cannot make a dipole bond."

Upon asking what would happen if they added too much fluorine atoms (a CF_3-group, for example), they could after some thinking come up with the idea that it could also be too big and not fit anymore.

> Group C: "Bigger... Oh, it won't fit anymore. So when you create a very long chain, it could also not fit anymore due to the form?"

Group D really focused on the chain length and the hydrophobic / Vanderwaals bonds that could be stronger. In the second measurement component, they were able to identify electronegative interactions and the chain length, causing hydrophobic interactions.

> Group D: "The R-group in analog 1 is longer, so more hydrophobic. It can make stronger Vanderwaals bonds. The F-H bond is more electronegative [referring to the dipole bond between two molecules]."

They were scored in level 2 mainly because they did not identify a lot of different interactions and because they made some errors in terminology, for example in which hydrogen atoms can actually form hydrogen bonds. However, they were able to draw their analog in the binding site and identify why they found it logical that a hydrophobic chain would face the benzene ring and

they did correctly identify the hydrogen bonds in their drawing. To summarize, they showed some characteristics of either level 1, 2 and 3 and in the end mostly fitted in level 2.

Conclusion and discussion

Conclusion

The aim of this research was to add knowledge about how students reason about molecular interactions and to what extent this curriculum unit helps them to improve this reasoning. We aimed to answer this research question: To what extent do students improve in their reasoning on molecular bonding/interactions during a context-based curriculum unit with a molecular modeling approach? From the obtained results, we can conclude that students improve in their reasoning during the implementation of the curriculum unit. In the measuring instrument, every group improved at least one level. This indicates that they learned how to reason about various interactions and are also to some extent able to connect interactions to characteristic groups in the drug molecule and enzyme.

Implications

Modeling skills. As has been mentioned in the theoretical section, it is important to build modeling skills with students: understanding of spatial molecular structures and the ability to transfer between molecular representations and chemistry understanding levels (Dori and Kaberman, 2012). This unit contributes to that by connecting microscopic properties such as molecular interactions to macroscopic properties, such as the effectiveness of a drug against a disease. The results furthermore confirm results of previous studies on computerized modeling: that the use of virtual learning environments improve the modeling skills of students (Dori and Kaberman, 2012). The use of an ICT-tool like ArgusLab is therefore useful, as this is a vizualisation tool that creates 3D models and shows the interactions between molecules. A suggestion for future research would be to redesign this unit for use without the ICT-tool in order to measure to what extent the ICT-tool itself contributes to student' understanding.

Drug development. The use of the drug development context contributes to motivation of students, as has been stated before by Bode (2016). Students learn how abstract chemical knowledge has implications in the real-world. Further research can be conducted to find out whether this can also be achieved using other contexts, such as polymer or environmental chemistry.

Limitations

Context. Even though the results indicate that the drug development context is useful, it can also be considered a limitation, due to its complicatedness. The chemical theory behind drug development exceeds the high school curriculum at some points (for example, several interactions such as pi-pi stacking are not covered). Experts also mentioned that the choice for malaria disease makes it complicated, because the mutations in the target enzyme are not directly visible in the binding site, which makes it hard for students to see and reason about changes in binding energy. It could be useful to search for another target within the drug development context.

ICT-tool. The use of the ICT-tool ArgusLab also results in some limitations. The program worked well overall. However, ArgusLab is not able to actually put a surface over the protein, while this was an important step that one of the experts named. Also, it is a very old program, and it could still be worth wile to search for a newer program with the same characteristics, which is still easy to use. At the school where the unit was tested, it was not possible to actually install the software on the normal laptops, it had to be done on some old laptops. This resulted in a poor (slow) performance sometimes, which does not improve the motivation of students. Improving the effectiveness of the ICT-tool might not only increase motivation, but also increase the speed by which students are able to complete the assignments.

Practical limitations. Furthermore, there were some practical limitations in this research. The study turned out to be a small scale pilot, because there were only 4 groups that actually did all the measuring components. There was a lot of shifts in presence in class due to an extracurricular project that allowed several students to be absent sometimes. Only three weeks were available and some things took some more time than anticipated. This can also be related to the speed of ArgusLab. Therefore, we had to skip some activities that could have helped with the understanding of the students.

Implications for teaching

The results from Bode et al (2016) already showed that this unit shows potential for actual implementation in the classroom. This study adds to this by stating that, next to being motivating and interesting to students, the students also gain understanding and improve in their reasoning about molecular interactions. It is therefore highly recommended to further improve and extend this unit in order to build modeling skills with students and therefore help them to better grasp chemical (and scientific) concepts.

Reference list

Bode, J. (2016). Integration of 3D Visualization and Molecular Interactions Using Molecular Modeling within the Context of Drug Development Against Malaria Disease for Learning in Chemistry Education. (Master thesis). Retrieved from Utrecht University Repository.

DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, *37*(6), 582-601.

Dori, Y. J., & Barak, M. (2001). Virtual and physical molecular modeling: Fostering model perception and spatial understanding. *Journal of Educational Technology & Society*, *4*(1), 61-74.

Dori, Y. J., & Kaberman, Z. (2012). Assessing high school chemistry students' modeling sub-skills in a computerized molecular modeling learning environment. *Instructional Science*, *40*(1), 69-91.

Eilks, I. Teaching 'Biodiesel': A sociocritical and problemoriented approach to chemistry teachers and students' first views on it. *Chemical Education Research and Practice* 2002*, 3*, 77-85.

Giere, R. N. (1988). *Explaining science: A cognitive approach*. University of Chicago Press.

Gilbert, J. K. (2005). Visualization: A metacognitive skill in science and science education. In *Visualization in science education* (pp. 9-27). Springer, Dordrecht.

Gobert, J. D., & Pallant, A. (2004). Fostering students' epistemologies of models via authentic model-based tasks. *Journal of Science education and Technology*, *13*(1), 7-22.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, *28*(9), 799-822.

Harrison, A. G., & Treagust, D. F. (2000). Learning about atoms, molecules, and chemical bonds: A case study of multiple-model use in grade 11 chemistry. *Science Education*, *84*(3), 352-381.

Holbrook, J., & Rannikmae, M. (2007). The nature of science education for enhancing scientific literacy. *International Journal of Science Education*, *29*(11), 1347-1362.

Hurd, P. D. (1998). Scientific literacy: New minds for a changing world. *Science Education*, *82*(3), 407-416.

Jorgensen, W. L. (2009). Efficient drug lead discovery and optimization. *Accounts of Chemical Research*, *42*(6), 724-733.

Juntunen, M. K.; Aksela, M. K. Improving students' argumentation skills through a product life-cycle analysis project in chemistry education. *Chemical Education Research and Practice* 2014*, 15*, 639-649.

Kolstø, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science education*, *85*(3), 291-310.

Kroemer, R. T. (2007). Structure-based drug design: docking and scoring. *Current Protein & Peptide Science*, *8*(4), 312-328.

Laugksch, R. C. (2000). Scientific literacy: A conceptual overview. *Science Education*, *84*(1), 71-94.

Peterson, D. S., Walliker, D., & Wellems, T. E. (1988). Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in falciparum malaria. *Proceedings of the National Academy of Sciences*, *85*(23), 9114-9118.

Prins, G. T., Bulte, A. M., van Driel, J. H., & Pilot, A. (2008). Selection of authentic modeling practices as contexts for chemistry education. *International Journal of Science Education*, *30*(14), 1867-1890.

Roberts, D. A., & Bybee, R. W. (2014). Scientific Literacy, Science Literacy, and Science Education. In *Handbook of Research on Science Education, Volume II* (pp. 559-572). Routledge.

Sadler, T. D., & Zeidler, D. L. (2009). Scientific literacy, PISA, and socioscientific discourse: Assessment for progressive aims of science education. *Journal of Research in Science Teaching*, *46*(8), 909-921.

Sanger, M. J., Brecheisen, D. M., & Hynek, B. M. (2001). Can computer animations affect college biology students' conceptions about diffusion & osmosis?. *The American Biology Teacher*, *63*(2), 104-109.

Schnotz, W., & Kürschner, C. (2008). External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. *Instructional Science*, *36*(3), 175-190.

Appendix A

Curriculum Unit

The unit consists of five chapters. The first chapter is an introductory chapter where students get familiar with malaria and drug design. It is mostly a theoretical description of how the malaria parasite works. Three ways to design a drug are highlighted: whole-cell screening, target-based screening and lead optimization. Students are asked to do an assignment in order to find out why it is necessary to find new drugs against malaria.

In the second chapter, chemical theory on proteins, their structure and amino acids is explained and it is also made clear that the target in this unit will be the DHFR protein, responsible for the synthesis of parasite DNA. Different mechanisms of enzymes are mentioned, like competitive and non-competitive inhibition. Students are familiarized with the Protein Data Bank. Next, the drug molecule (pyrimethamine) is shown and some theoretical concepts, such as binding energy, are explained.

In the third chapter, the students will start the molecular docking, using ArgusLab for the first time. In this chapter, an instruction manual is mentioned where every step the students need to take is written down. Through assignments, students learn how to use ArgusLab and start to think about alternatives for pyrimethamine. This continues in the fourth chapter, where they also get the final assignment: to design and test their own alternative to pyrimethamine.

Appendix B

Lesson scheme for the try-out (Dutch)

DATUM/LESUUR	LES	INHOUD LES	BENODIGDHEDEN
WOENSDAG 7	1	• Introductie module/opzet (Laura) • Uitleg/doornemen hoofdstuk 1 (Corina): drie manieren van geneesmiddelen ontwikkelen • **Maken opdr 1 t/m 3**	• 6 laptops met internet • Module H1+H2 geprint
VRIJDAG 6/7	2/3	• Afmaken opdracht 3 • Uitleg eiwitten • Introductie Arguslab • **Maken opdr 4**	• 6 laptops met arguslab • Arguslab instructie geprint
WOENSDAG 1/7	4/5	• **Maken *Pretestopdracht op antwoordvel*** • Introductie docking (H3) • Opdrachten doen met docking • **Maken opdracht 5 t/m 8**	• Antwoordvellen voor **pretest** • 6 laptops met arguslab • Module H3+H4 geprint
VRIJDAG 6/7	6/7	• Werken aan opdracht 5 t/m 8 • Start hoofdstuk 4 • Docking opdrachten/ontwerp	• 6 laptops met arguslab • Antwoordvellen voor **meetinstrumentvragen**
WOENSDAG 1/7	8/9	• Werken aan dockingopdrachten • Interviews in groepjes (Laura) 3x het 1e uur, 3x het 7e uur	• 6 laptops met arguslab • Antwoordvellen evt. • Hardopdenkprotocol voor interviews

Appendix C

Interview protocol expert interviews

Protocol

General

1. Are there any questions about the material I sent you? Do you understand what the goal of my research is?
2. Which steps do you take when performing lead optimization - what does that process look like globally?

■ Note the various steps in order to continue asking about those questions

Specific

3. Zooming in on the different steps.
    a. What is important about this step / argument?
    b. When applying this to the case of malaria - if you were to find an alternative for pyrimethamine - what would you do and what would be the result of this step?
4. Do you have any cases in your current/recent research that are comparable with the malaria case (lead optimization)?
    a. If yes, what does that process look like, what are the steps being taken there?

Finishing

5. In the case of using a docking program, which program do you use? Is it paid or freeware?
6. Do you have any comments or questions concerning this interview? Is there anything that has not been asked that you would like to add?

Thank you so much for agreeing to this interview!

Appendix D

Think aloud protocol

1. What are you doing right now, what is your plan?
2. What is the reason you are changing this?
3. What do you expect from exchanging this characteristic group?
4. Can you explain these results?