Al-Ishlah: Jurnal Pendidikan
Vol.xx, x (xxxxxx), pp. xx-xx
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v14i1.973

Database and Molecular Visualisation for Learning Media of Teaching Chemistry

ARTICLE INFO

Keywords:

molecular;
modelling;
computational;

Article history:

Received xxxx-xx-xx
Revised xxxx-xx-xx
Accepted xxxx-xx-xx

ABSTRACT

We measure student-learning outcomes during computational courses using molecular visualization software and databases in the chemistry and pharmacy. The instruments of this research are the results / answers of final exams and the assignments of the students in the semesters of 2018/2019 and 2019/2020. A hundred students enrolled in these courses per year. The study was conducted for two years. We analyze the results, six assignments for practical and four essay questions for theoretical tests. Students were asked to visualise molecules using the requested application in practical test. The assesment shows that project scores with the topic of using databases and visualization, a 'good' score profile of the results the two years. The student has 'good' score on small molecule database and visualisation. The students mastered the Avogadro application more than VMD. Students generally understand to get data (computer files) from the database and then the visual constructions requested by the questions.

Corresponding Author:
Nirwan Syarif
Universitas Sriwijaya, Indralaya, Indonesia; nsyarif@unsri.ac.id

1 INTRODUCTION

The spirit of education 4.0 is to create an "independent" learning climate and provide space for students to be independent and independent in Learning. The advantages of independent learning include improving academic performance; increasing motivation and self-confidence; increasing students' awareness of their limitations and abilities makes it easy to manage their potential and make teachers aware of each student's ability to give different assignments and build. Today's world of education is closely related to information and communication (Fetaji et al., 2018). Information technology has transformed into something very attached to the world of education in today's modern era. Starting from computers, the internet, and other hardware and software, they are continuously being improved so that their utilisation can be perfect in the use of information technology as an

educational media, such as telephone, computer, internet, e-mail, social media and databases (Fitrani Dinda Fadhilah et al., 2021).

Media that supports the development of teaching technology is needed in supporting knowledge to understand concepts and gain the skills especially in chemistry (Gavezzotti & Presti, 2016). Non-physical learning materials such as chemical databases can describe chemical molecules that number in the millions or even billions. Some of these molecules are natural molecules that function as life support. It is contextual to deliver it in the class. This learning material is also very factual because everything is described based on daily experiences. This factual depiction of chemical-biological molecules is one of the characteristics of molecular modeling. Molecular visualization is built on the ideas of various experts (Kozlíková et al., 2017), such as chemists, biologists, physicists, mathematicians, computers, electronics, and even philosophers. The vision for molecules are the wonders of the thought before computers and measuring instruments were developed. Advances in computers, where computers can work quickly to process images of molecules composed of thousands of atoms with atomic orientation in three-dimension space, computer monitors are precisely matched by increasingly microscopes that can see down to the atomic level. Chemistry / chemical databases are now available online and free on internet with a wide variety. Databases containing molecular structures such as the Protein Data Bank (PDB), ZINC, Crystallograpic Online Database (COD) have been used in computational courses for many years, as well as the softwares, such as VMD and Avogadro (Rayan & Rayan, 2017).

Databases in education support the idea of "Freedom of Learning" because they are independent, contextual and innovative. In the other hand, visualisation in chemistry is part of the general multimedia explosion in science education, which is characterized by the design of custom visualizations to illustrate particular structures or concepts. Computer-supported databases have become valuable tools for several research laboratories in the industry and academia (Higgins et al., 2022). Applications made separately based on specific needs, are combined to provide ability to visualize molecules so that they can be scientifically accepted by students, teachers, researchers, and practitioners

Computational chemistry / pharmacy is an important subject in chemistry teaching and learning activities, this is because computational chemistry has practical values in the form of the ability or skill to make concrete abstract concepts, bring objects that are dangerous or difficult to obtain in the learning environment, can calculate complex molecular properties, can be used as a visualization and animation tool, can also calculate molecular properties and their changes as well as perform simulations on large systems and apply the program to real chemical systems.

One of the most important topics in computational chemistry is molecular visualization, because it is the interface between humans and computers that communicates several other concepts used in computational courses. The results of observations on the final grades of students who attend lectures are still unsatisfactory, so it is necessary to look for the causal factors of this problem. This is due to the students' lack of understanding and systematic approach of complex chemical structures and representation of molecules (Karonen et al., 2021). One attempt to overcome this problem is to develop software that describes molecules in three dimensions. Today's computer visualization applications have even become interactive and manipulative. Users can use various representations highlighting different structural features and zoom, pan, and rotate molecular images. This ability then makes

The visualization has been help student in studying protein molecules and understand course material more deeply in the way of improving their intuitive thinking (Argüello & Dempski, 2020). Erica, et.al (2022) reported students improvement on crystalography understanding in solid state chemistry course based on molecular visualization. Maia, et.al (2016) have shown that students' ability to visualize the enantiomer of a simple molecule in three dimensions improves their skill in organic chemistry. In a series of highly detailed studies, Burgin et al. (2018) found significant gains in high school students' understanding of simple molecular representations after using molecular visualization software. In the study described in this paper, we examined changes in students' concepts of a more

complex molecule, a protein, resulting from a relatively short exposure to molecular visualization. We also examine the effects of expanding various representations / styles of molecular structure originated from the database as part of their project. To get a definite picture of how students understanding about molecular visualization and databases, an analysis of the student's work is carried out, namely test scores and assignment scores.

2 METHODS

The activities on descriptive qualitative research for the students were conducted in chemistry / pharmacy department of Universitas Sriwijaya. This study analyzes and describes data in the form of numbers, namely test scores and assignment scores which are then used to explain computational chemistry / pharmacy learning outcomes, especially on the topic of molecular visualization and databases. The data were the results/answers of the assignments in the project and the final test of chemistry and pharmacy students in the semesters of 2018/2019 and 2019/2020 in taking computational class. The number of respondents in this study were 200 students. The results describe students' understanding of the theme of the lecture, namely for database and molecular visualization.

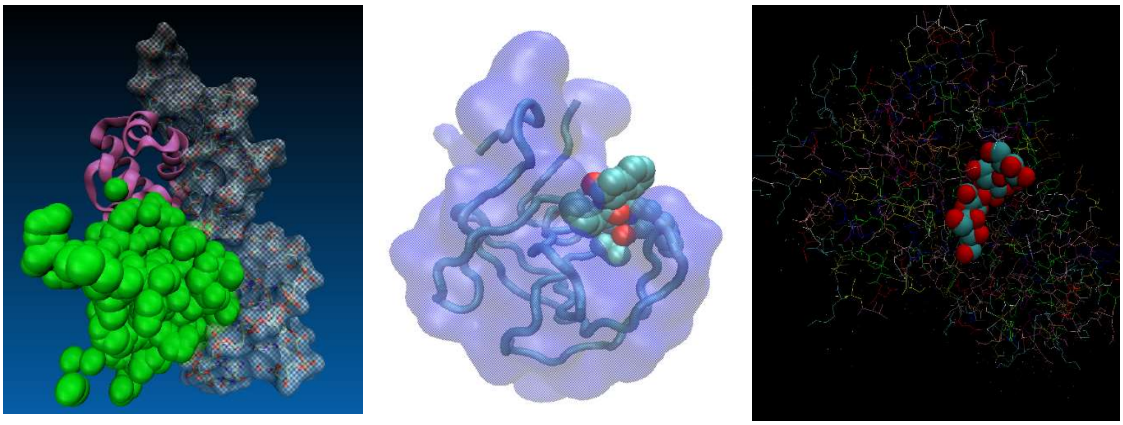

Figure 1. Visual reference of the downloaded file from the PDB and was completed using VMD.

The results of final semester exams and assignments/projects in computational chemistry/pharmacy courses were used. The project consists of two parts namely (1) using online databases, namely protein data bank (rcsb.org), ZINC20 (zinc20.docking.org) and COD (crystallography.net). Students then download the file containing the structure and (2) visualize the molecular structure / molecular system adapted to the drawing command. The test consisted four essay questions that can be used to provide an overview of the use of databases and computer visualization applications. The questions are (1) what is meant by modeling. What does this have to do with molecular visualization? (2) what kind of molecular model can be visualized? (3) Describe molecular databases and their application?; (4) What is the role of molecular visualization in databases?

The students' practical understanding of databases and molecular visualization is assessed from the lecture project. The task in the project is to download three files from PDB (rcsb.org) and two files from ZINC20 (zinc.docking.org) and one file from COD (crystallography.net). The three PDB files with extension pdb, are 3cro, 1hsg, 1fqa respectively in Figure 1 (A), (B) and (C).

Figure 2 (D), (E) and (F) are ZINC169366333 (Porphyrin), ZINC901155 (Galactose) and 5 water molecules, and 1000063 CIF (Cu2O). All files are provided with screenshots of the application that are used as references for student work. The files are opened using VMD. Avogadro is given the available styles for the file components so that the appearance / visualization of the molecule is the same as the reference image (Figures 1 and 2). Full marks are given if all the visual components of the file are met, counting from 6 questions.

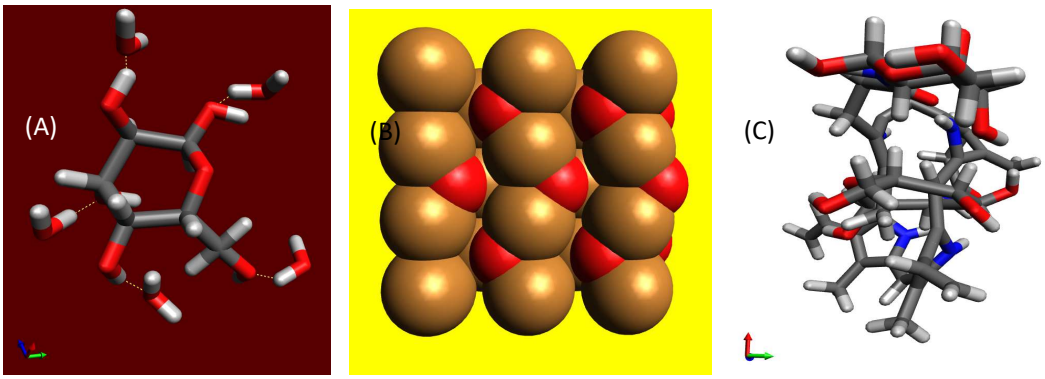

Figure 2. Visualizing the references of the downloaded file from ZINC20 and COD performed with the Avogadro

Students' theoretical understanding is assessed from the answers to the final semester exam. We categorize students' answers into three, namely: 'good', 'fair' and 'low'. The 'low' category is given if the student gives an answer, others, I don't know or it is empty. The results of students' answers to the four questions are discussed one by one as sub-chapters. The calculation of test scores is carried out as in large molecules. Reference images are given as a guide in doing the task. The image in Figure 1 and 2 show the references that students were expected to be able to do using the VMD and Avogadro. The value of each number is calculated from the components of practice that can be done by students and the total value for practice is the sum of the values of each number divided by 6. Calculation of practice is calculated using the formula.

value per file = visual component worked on/ (total visual component + database) x 100%
average value = sigma (value per file) / 6

value calculation results. The value for the incomplete theory is calculated by the formula, where the total question is 4

value for 1 question = fulfilled components A/ total visual questions x 100%
value for 4 questions = total % total score / 4

3 FINDINGS AND DISCUSSION

This study focuses on two problems: first, how was the students' understanding of molecular visualization, especially its definition, types, and uses; second, how was the students' knowledge of the database, its usefulness, and its relation to molecular visualization, theoretical and practically. We examined the effect of the variables, i.e the knowledge and time of course enrollment on students' level of understanding. Detailed data from students' answers in both semesters, pratically and theoreticallt are presented below.

Practical Tests

This course taught offline were conducted in a computer laboratory with 50 students facilitated by computers, networks, and the internet. Lectures and labwork were integrated, where we delivered theory before the labwork was carried out. Access to the database can be done directly from the available browser because the databases used are open-access, namely PDB, ZINC, and COD. Each student is guided to download the applications used, namely VMD and Avogadro, for molecular visualization tools. The teaching is carried out after the application is installed on the computer used by students. The project consists of several steps, namely: searching for files in the database, downloading files, visualizing the molecular structure on a local computer, and styling parts of the molecule to match the given molecular model.

A visualisation is an approach that uses multiple representations in developing visuospatial thinking in chemistry (Kiernan et al., 2021), especially for the student. The teachers are expected to be able to use molecular visualization tools as well as using physical molecular kits. This visualization application provides an interactive way to view and dynamically manipulate multiple representations of molecular structures. The student can also use this application in other courses to adjust the existing curriculum and activities. For this reason, we assess the assignments of students taking part in computational chemistry/pharmacy lectures to work on the final project. We also give the assessment results in three categories: 'good', 'fair', and 'low'.

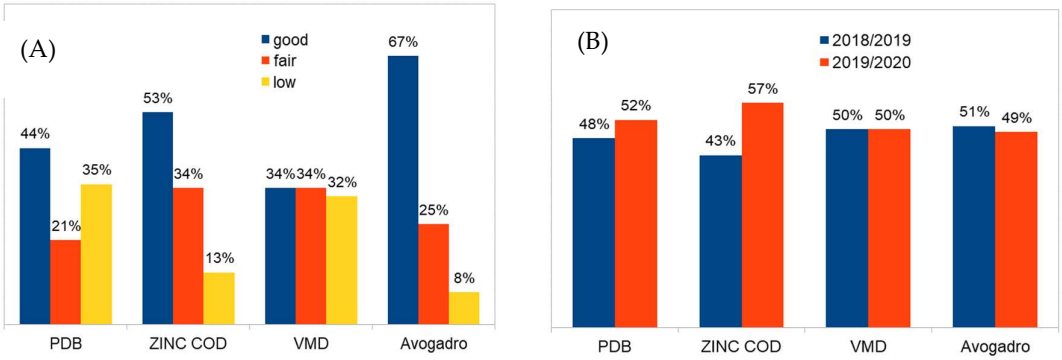

Figure 3. (A) The results of data processing for the tests with the topic of using databases and visualizations, (B) a 'good' score profile for college exams for two years

Students were asked to represent molecules using the requested application to measure students perceptions of the usefulness of visualization software. The project is considered 100% complete if the presentation made follows the reference image provided. Figure 3 shows (A) the assessment results on project scores with the topic of using databases and visualization, (B) a 'good' score profile of the results of the two years. We can see from the plot (Figure 3A) that 53% and 67% of student has a 'good' score on ZINC COD database and Avogadro visualisation. These scores have not changed much for two years (Figure 3B). The profiles for VMD and PDB are relatively lower, indicating that students are more familiar with small molecule databases with more straightforward visualization.

The interface on the search page and results look more straightforward and intuitive, showing that mastery of the application is followed by knowledge of the database. Databases for small molecules, such as those of ZINC20 and COD, do not use the complex interfaces of VMD. We found that asking students to use a ZINC20 or COD database and then displaying the crystal structure required fewer additional instructions from written instructions as project implementation instructions compared to what was done with PDB. Avogadro's interface looks simpler than VMD's. The features in the VMD interface are far more numerous with more diverse explanations, so it is not easy to understand. The students only need to master the application's practical materials to understand basic database concepts and molecular visualisations.

Students mastered the Avogadro application more than VMD. Students generally understand how to get data (computer files) from a database and then the visual constructions requested by the questions. Molecular structure construction involves building a physical model from a crystal structure model kit (Simmons et al., 2016). The instructions also guide students to use a web-based unit cell visualization tool and intersperse the activity with short segments of hands-on instruction using several drawn unit cell representations.

What is meant by modeling?

Figure 4 shows the assessment results of the theoretical test. The questions related to the databases and visualization were conducted in essays. We can see in Figure 4A that the students have a 'fair' answer on the definition of modeling and databases. Exact solutions were obtained from students answering more or less the same as what was described previously, accompanied by examples of images that visualize molecular models both in 2D and 3D. This question then relates to the type of molecular model displayed in software such as Avogadro and VMD. The results of data processing showed that quite a lot of students could not answer the questions. Most of them only provide general answers to molecular models that are often exemplified in lectures for both 2D and 3D. As previously explained, molecular models are not only for visualizing molecules but can also be used to describe specific molecule characteristics. Structural characters of molecules such as enzymes can be explained not only by using this general visualization but by using cartoon visualization. Visualization with cartoon style for enzymes makes it easier for students to determine the dominant structure of enzymes. Students can also understand the concept of extended structure in crystals, determine the formula for ionic compounds from the unit cell, and determine coordination numbers.

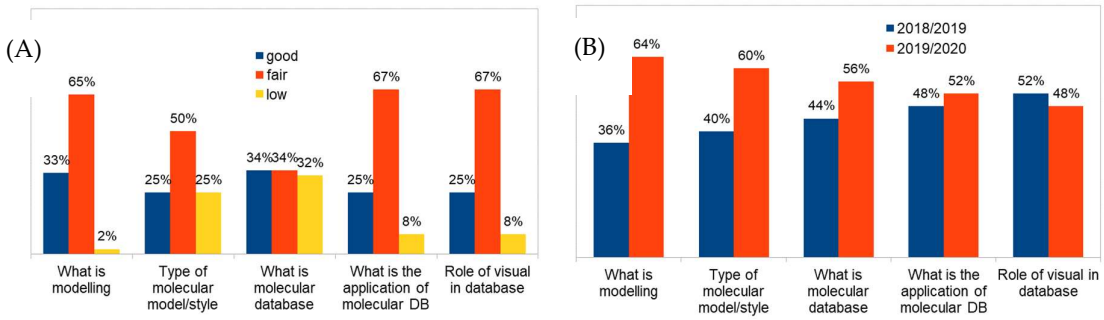

Figure 4. A) The results of data processing of practical assignments on the topic of using databases and visualization, (B) a 'good' score profile for college exams for two years

When the complete answer to the question What is meant by modeling? What does this have to do with molecular visualization? The most frequent answer given by students in the two years was ‘fair’, 65%, and the 'good' answer, 33%. For student results in 2018, the 'good' answer is around 36 %, while in 2019 is 64%. In particular, the emphasis on 'fair' answers in 2019 was reflected in the complexity and diversity of student answers, as measured by the average number of categories in each student's responses, which increased significantly by 32%.

As an exact answer to the question above that, modeling is a way to represent or describe or explain an object, actual situation (reality), process, system, or concept, simplified by only paying attention to the essential and ignoring factors that are deemed unimportant in the study carried out or idealization. Molecular modeling is an idealization of chemical compound molecules. It took more than a hundred years to describe the molecule. The molecule of a chemical compound itself was originally an abstract entity. Most of the students gave precise answers regarding the definition of molecular

structures. Students' responses are interfered with by the definition of mathematical models, which are also studied in other courses, such as mathematical models and physical models. Mathematical models, as used in teaching kinetics, are formulas that contain numbers and symbols so that they become equations and inequalities. While the physical model is photos, pictures, paintings, and others that are physical.

What kind of molecular model can be visualised?

Conceptualization of complex macromolecular structures motivates scientists to simplify computer graphic images representing specific entities. The computer can visualize molecules using computers at several levels, namely the quantum level (electrons), molecules, and molecular groups. Visualization of molecular structures, in some ways, shows particular characteristics that are physically more obvious than atomistic representations. In other words, the visual levels complement each other depending on the details of the character to be explained (xxxx). With the growing understanding of biological structures, the need for new visualizations. Molecular visualisation supports materials' knowledge, characteristics, and interactions by creating molecular structures. The benefit is that it helps the 'rational' design of new molecules, for example, for catalysts or drugs. Visualising biological molecules, such as enzymes, focuses on graphical depictions of these biomolecules' structure, interactions, and functions. The complexity of the biomolecular and all functional biological units that occur in biological cells increases when other cell components or molecules, such as cell walls and DNA, are considered.

Learning key patterns from complex observations is practical to use in visualising molecules. This ability is developed from the bond character, bond form, bond distance, bond angle, and the molecular plane of a molecule. Their values predict its interaction with large molecules and relate it to its biological function, such as hydrophobicity/lipophilicity, three-dimensional geometry, surface electric charge, molecular contours, reaction sites, polarity, and others. These characters are a treasure trove of knowledge that has just been discovered with the existence of today's technology. This descriptor is in the scope of research for cheminformatic applications close to drug discovery. The student can use it within the range of teaching to explain the relationship between molecular structure and physical phenomena or even experimental measurement results.

Selecting a particular representation relies heavily on the descriptor set for good predictive performance. The existing descriptors each encode different information, and no single model works well, so multiple descriptors are needed to link to a single molecular structure. For example, molecular weight or polarity can correlate well with other properties, such as boiling point, but are not very precise when associated with complexly bonded molecules. Thus, most molecular representations encode information optimized for a particular use.

Analysis of the electronic structure of a molecule by calculating the ground state, excited state, and transition state that occurs during chemical reactions can be derived from understanding the electron character of the atoms that make up the molecule. The resulting data, which is an electronic structure analysis, is in the form of the expected value of the density or flux function or of the electron-nuclear wave function. The formula is the energy value as a geometry function (atomic coordinates). Energy values are also related to molecular orbitals. At this level, molecular visualization requires multi-plane depictions to help reveal the phenomena of physical processes, such as color and light emission. Visual analysis of electron density topology provides information about the spatial domains associated with individual atoms.

This question demands more knowledge of the scope of molecular visualization. The results of processing test score data for the question "what kind of molecular model can be visualized?" shows that the answer most often given by students in two academic years is "fair." For student results in 2018, the 'fair' answer is 40%, while in 2019 is 60%.

What is molecular database?

Accessing chemical information stored in various databases using computers and the internet has increased significantly, especially in the era of big data (Bird & Frey, 2016). The database's dimension is expanding in terms of structure and capacity. Sophisticated search algorithms run in databases are essential apart from molecular visual applications. Databases supported by computer devices are helpful tools for research and development in industry, academia, and teaching. A database in the field of chemistry has details about chemical reactions and chemical compounds such as toxicological, spectral, physical, or chemical characterizations with a search facility. The most common for information about a chemical compound or its chemical products are searches by chemical name, IUPAC, empirical formula, structure (Heller et al., 2015), or substructure, and even by using journals reporting on the compound/molecule in question (Llanos et al., 2019). Chemical databases are an essential library resource for discovering drugs, catalysts, membranes, polymers, bullet-resistant fibers, adsorbent composites, and other modern materials.

Figure 5. PDB main page on rcsb.org; accessed October 7, 2022

Data in Figure 4 on the question "What is the molecular database?" shows that all answer categories are relatively the same" For student results in 2018, the 'good' answer is 44%, while in 2019 is 56%. The explanation for the molecular database is relatively hard to describe to the student. This profile is

caused by the broad scope or range of the database. The databases in chemistry are very diverse for such educational resources supporting molecular exploarations (Zardecki et al., 2022) .

Databases for large molecules, such as PDB (rcsb.org) (Figure 5), provide a place to store molecules such as enzymes and DNA. RNA and biomembranes. Enzymes are hosts as molecular targets for ligands other than natural substrates. There are known in-vitro and in-vivo stages within the scope of drug or medical research or the use of natural ingredients. With the use of computers as part of drug discovery, the steps are no longer two but three, with the addition of in silico (Edwards et al., 2021). Currently, some of the material inventions more or less pass through the computer design stage. The search on the database is carried out based on potential requirements for chemical, physical and biological activity, namely by identifying compounds/molecules that are suitable to be continued in the in silico research stage.

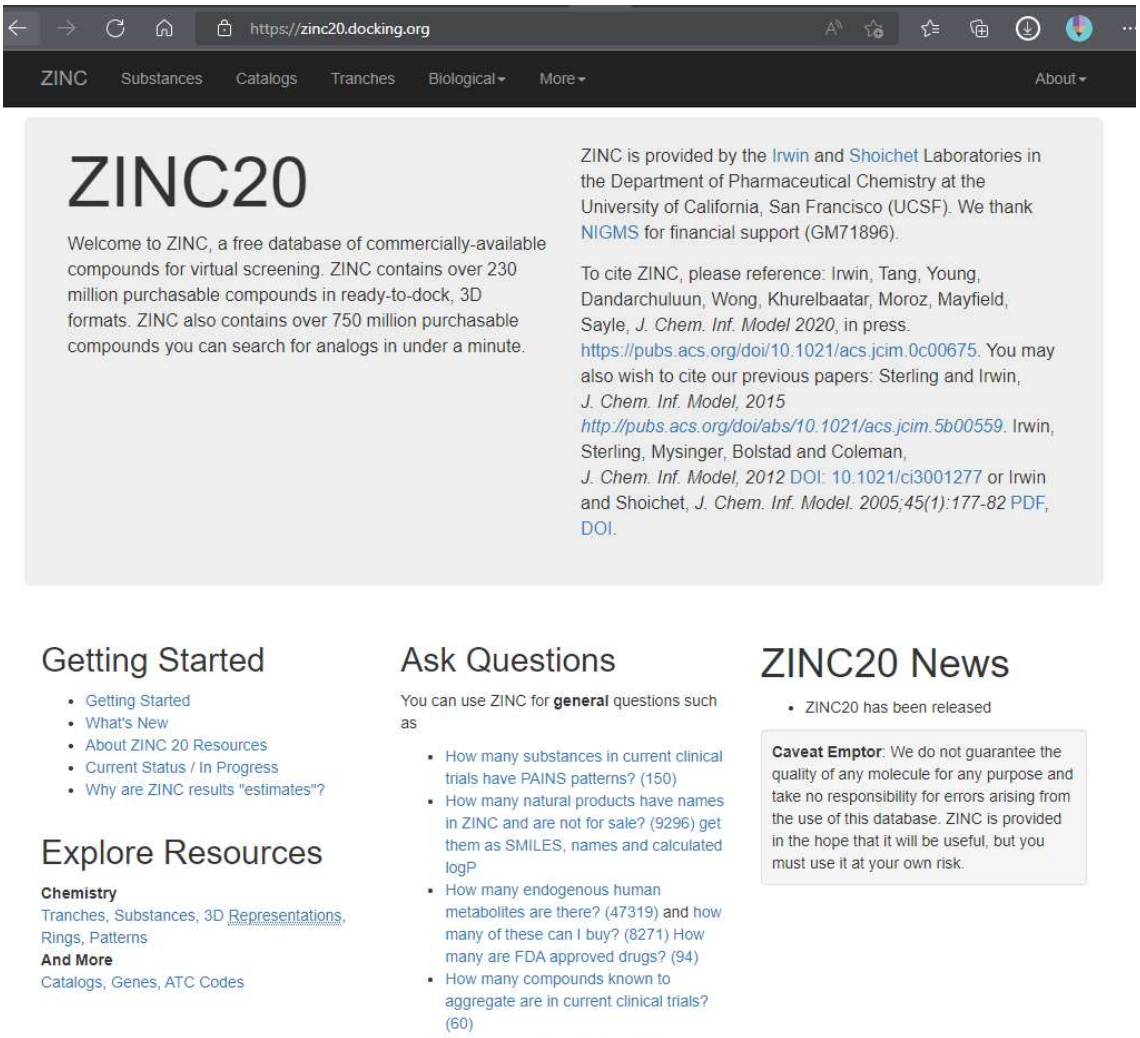

Figure 6. ZINC main page currently version 20 at zinc20.docking.org; accessed October 7, 2022

Databases for small molecules such as ZINC20 in zinc20.docking.org (Figure 6) provide templates for designing ligands or drug molecules that are substrates for enzymes (Irwin et al., 2020). The molecular structure of ZINC20 is treated as a template for the design of other molecules. A hundreds of different molecules are formed from one template molecule, distinguished from their functional

groups. These of molecules are then observed in-silico with their parent molecules, enzymes. The interaction between the ligands and their substrates was assessed using several parameters to conclude that some of the best ligand structures were said to be 'prospective' drugs. These best molecular structures are then feasible to be synthesized and enter the next testing stage.

ZINC provides downloadable 2D and 3D versions and a website that enables fast molecular and analog searches (Figure 7). ZINC already holds data on about two billion compounds on its servers. Users use ZINC by users is handled an interface that is always dynamic from time to time to respond to advanced user needs, storage growth, and device developments. Thousands of users use the ZINC website every month, including researchers, developers, students, and teachers. The exchange of data between users and servers can be seen through the number of terabytes of data downloaded every week. The students can use the statistics of "Usage' via the "About" menu followed by "Usage" at the top of each ZINC page.

ZINC's original focus was primarily to serve data needs on molecular docking with an analog-by-catalog approach and molecular similarity principles. Many researchers find it helpful to search for analogs in real time, where "what if" questions about compound availability can be explored. This ability then developed into the identification of analogs of compounds with a structure-activity relationship approach.

Searching the database based on potential requirements for biological activity (Figure 7A) or chemical structures (Figure7B) makes it easy to identify specific compounds/groups of compounds. Accessing chemical information on ZINC by utilising computer facilities has become more significant. The dimensions of all databases in terms of structure or reaction proliferate yearly. Therefore, computer-supported databases have become valuable tools for several research laboratories in the industry and academia.

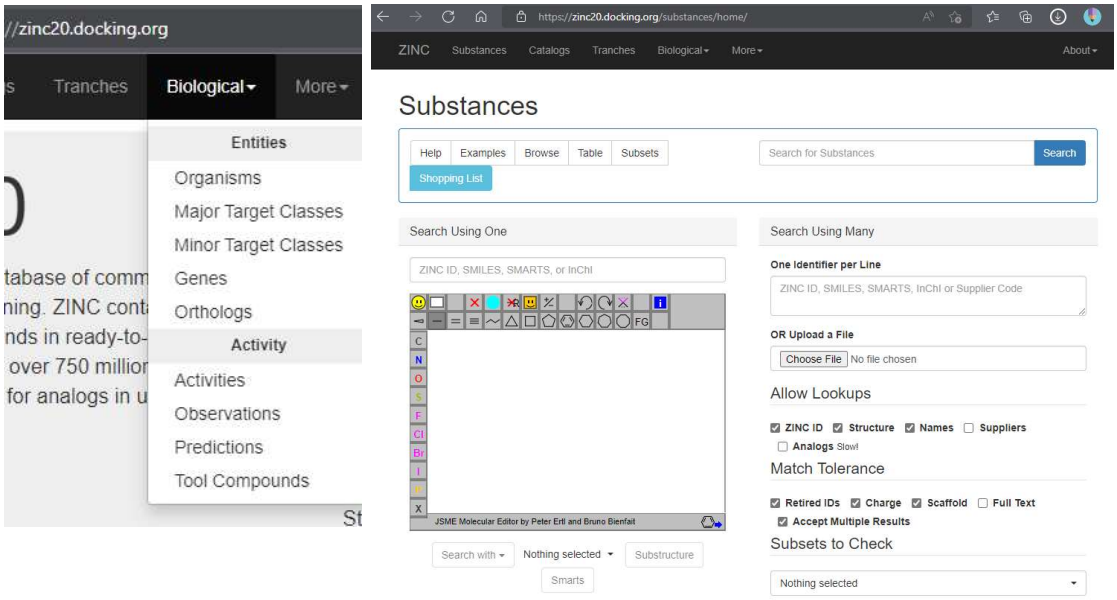

Figure 7. Search facilities available on ZINC

COD (crystallography open database) on www.crystallography.net (Figure 8) provides the coordinates of the atoms that make up the crystal structure of organic, inorganic, metal-organic, and mineral materials. The molecular structure stored in this database results from X-ray diffraction measurements of single crystals or powders, gamma photons, neutrons, electrons, or other particles.

We refined the results of the measurements geometrically with the results of calculations using the density functional theory (DFT) (Gražulis et al., 2009).

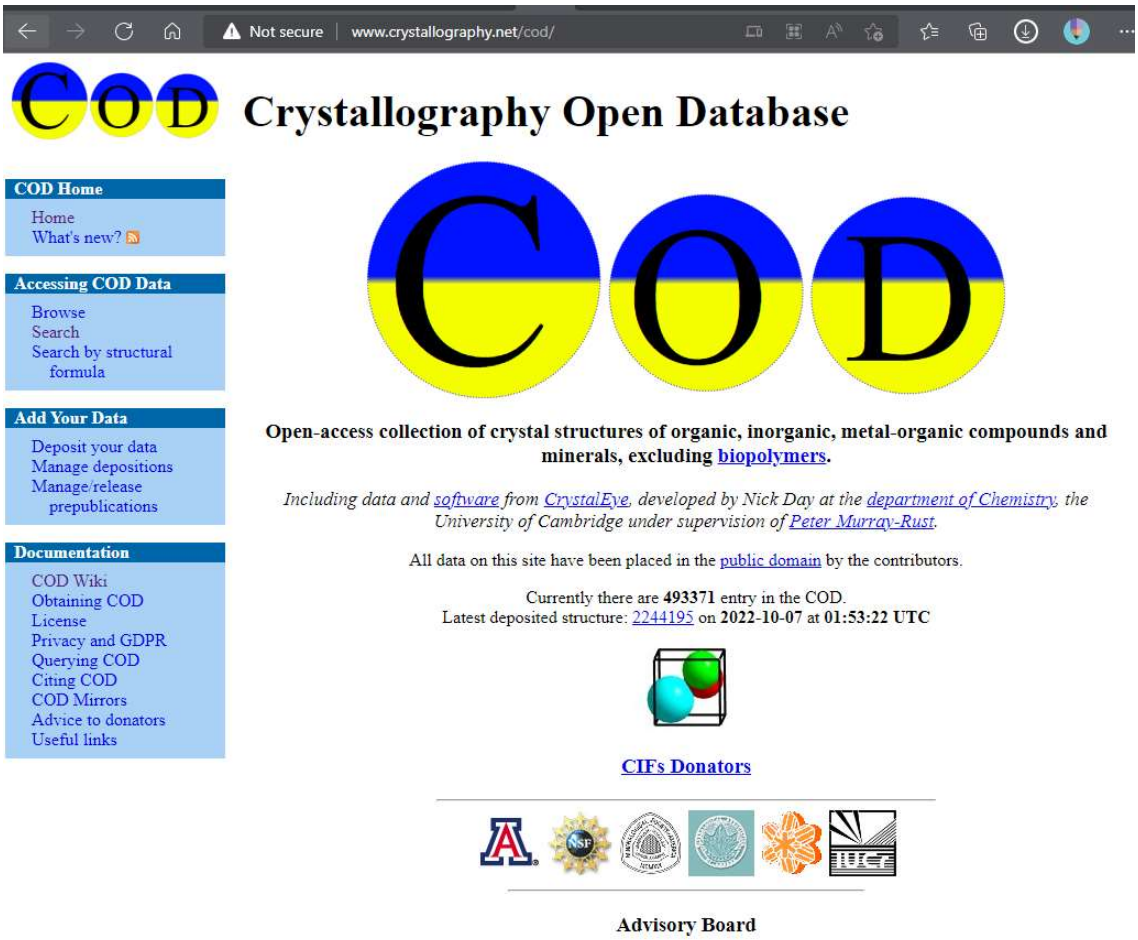

Figure 8. COD main page on crystallography.net; accessed October 7, 2022

Each structure deposited into a COD receives a unique seven-digit number, called a COD number. The COD number identifies a particular instance of the structure determination. As a rule, COD does not accept duplicate structures. If two structures of the same compound are published in two different peer-reviewed journals, they can be saved to COD and receive other COD numbers. Similarly, if the structure of some compounds is solved with much higher resolution or using more advanced methods, these new instances will be stored in the COD under the unique COD number. COD thus intends to capture and represent the current state of crystallographic knowledge. 'Old structures' can be filtered using crystallographic quality indicators, specific COD markers, unit cells, or publication dates.

What is the application of molecular database?

Each structure deposited into a COD receives a unique seven-digit number, called a COD number. The COD number identifies a particular instance of the structure determination. As a rule, COD does not accept duplicate structures. If two structures of the same compound are published in two different peer-reviewed journals, they can be saved to COD and receive other COD numbers. Similarly, if the structure of some compounds is solved with much higher resolution or using more advanced methods, these new instances will be stored in the COD under the unique COD number. COD thus intends to

capture and represent the current state of crystallographic knowledge. 'Old structures' can be filtered using crystallographic quality indicators, specific COD markers, unit cells, or publication dates.

The database contains information about the experimental conditions (temperature, pressure, or sample history) for each molecular structure. The unit cell is one of the input data when the user searches the database for comparison with the unit cell, which is also the identity of the molecular data stored in the database. Chemical formulas are checked for exact matches, and both formulas provided by users and calculated from the contents of the crystal structure (atomic coordinates, crystal density and molecular weight) are compared with each other (resulting in four comparisons if the formula provided by the author is in both files).

The range of compounds included in COD allows the use of COD for material identification using powder diffractograms and powder patterns calculated theoretically from COD entries. Even without 100% coverage of all known compounds (and elements), We can use COD to collect statistics on the chemical properties and structure of organic compounds, assuming that COD has a representative subset of the desired compound because COD can be accessed and downloaded for free by universities and students. COD can also be used for teaching, providing real-life and high-volume examples from scientific databases to formulate challenging data mining and processing tasks. The open nature of COD allows downloading this database or parts of it to any number of computers, thereby making use of many parallel resources available in grid/volunteer computing environments.

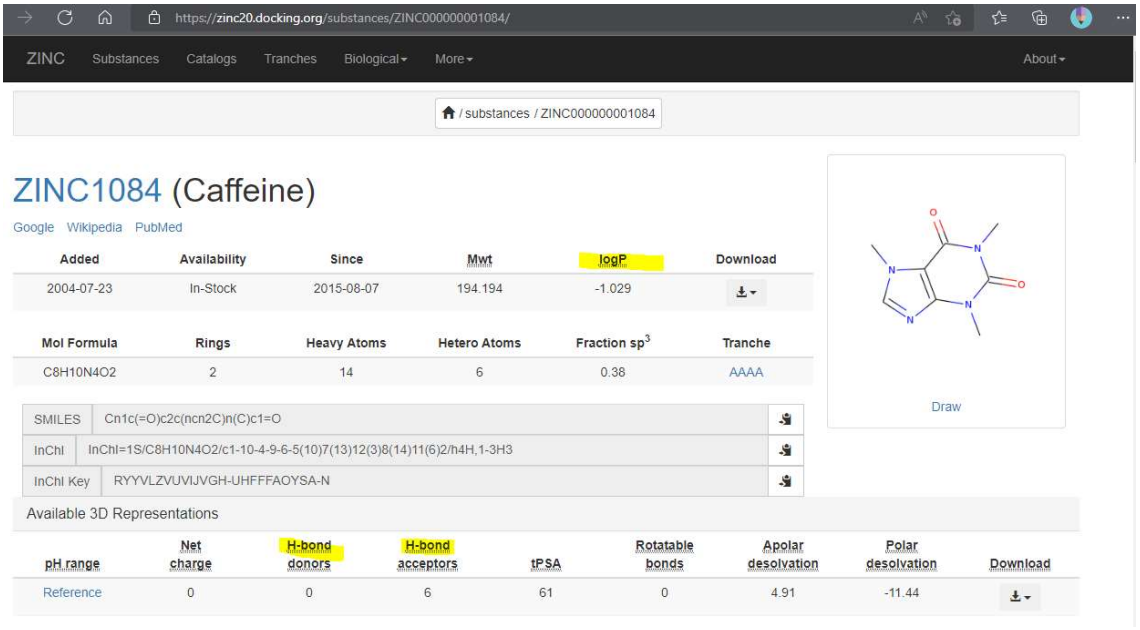

Figure 9. Some physical properties of the molecule that calculate on site, display in Zinc20 page

There are fundamental differences between the search results pages of the three databases, primarily related to the visualization displayed. The page on PDB displays a 3D molecular structure, ZINC20 displays a 2D structure, while COD does not display any structure at all. The difference in the appearance of the database is related to the application of the database, the value of the information contained in the database, and the effectiveness of the presentation of the data itself. 3-dimensional molecular visualization on PDB because the data displayed is macromolecule which will not be effective when displayed in 2-dimensional format. The presentation of the protein molecular structure on PDB precisely displays the tertiary structure of the protein, where with this visualization, some structural properties of the protein, such as hydrophobicity, and charge, can be seen immediately. The

structure of proteins determines their clinical significance. A simpler view is presented in the 2-dimensional visualization of relatively small molecules. Several molecular characteristics, such as functional group, molecular weight, and three conformational samplings three, were obtained from the 2D visualization. The available formulas, such as logP and hydrogen-bond acceptors, can easily calculate some quantities as shown in Figure 10. Even if it's displayed in 2D, it's necessary to set up a 3D view. Supporting data related to molecular structure, especially the formulation of 2-dimensional to 3-dimensional transformations can be used to build protonation models of molecules, tautomeric forms, the stereochemistry of racemic mixtures, and regioisomeric forms.

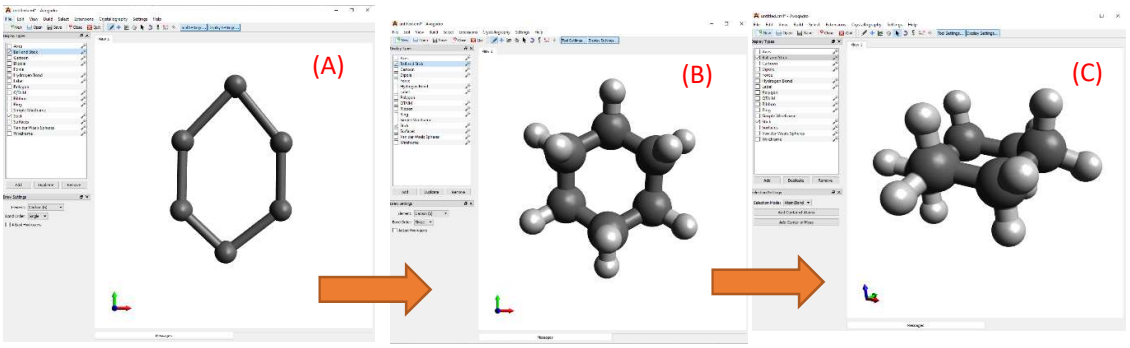

Figure 10. Visualization of the cyclo-hexane molecule, the complete depiction process and its conformation

What is the role of molecular visualisation in databases?

Visualisation in chemistry is part of the general multimedia explosion in science education, which is characterized by the design of custom visualizations to illustrate particular structures or concepts. However, the linkage between scientific concepts and visual representation is often not profoundly principled or algorithmic. General principles inform better illustrations of good visual representation and graphic design. In chemistry, however, a unified theory is used to make quantitative spatial predictions, which allows highly generalized rules to link theoretical concepts to visual displays. Expanded structural formulas that show the connectivity within a molecule by lines connecting the elemental symbols of the atoms are an example of effective visualisation in which a set of diagrammatic rules produces qualitatively informative diagrams of an unlimited range of molecules

The answer for the role of molecular visualisation in databases is a 'fair' given by students. Data in Figure 4 on the question "What is the role of molecular visualisation in databases?" shows that all answer categories in relatively the same" For student results in 2018, the 'good' answer is 52%, while in 2019, it is 48%. This data infers that the teacher provides more concrete examples of database applications so that students do not find it too difficult to provide examples for the application of molecular databases.

Crystallographic data combined with computational power drive the design of representations and the visualization of three-dimensional structures. The visualisation provides a general-purpose representation designed for molecule properties prediction. The development of visualisation prompted the re-optimization of the representation of molecular on accessing the databases. This feature creates new opportunities to utilise spatial information for molecular studies, such as quantitative struture – activity relationship and similarity analysis (Baskin & Zhokhova, 2019). Several accessible small molecule collections are available and provide start-up for the studies. Despite containing the three dimensional structure, there also provide data, such as tautomeric, protonate, and partial charging in molecules. Virtual filtering libraries for molecules can be constructed from the

combined information of 2D – 3D structures. Figure 10 shows the transformation process of a two dimesional design which is then generated automatically by the application to produce a three dimensional image. The visual variety of the database is associated with the application of the database (Hossain, 2019). In that case, the database application is not only used for research or development in the industry but also for education. Teachers use this application to encourage student involvement in understanding concepts in chemistry. The databases allow students to visualize and manipulate molecules. Therefore, it provides opportunities for students to perform computational measurements of physical quantities or to compare the results with experimental results in the laboratory.

4 CONCLUSION

The development of representations / styles in molecular visualization applications reflects research and technology trends as well as the information systems landscape in chemistry. Visualization in chemistry is part of the general multimedia explosion in science education, which is characterized by the design of custom visualizations to illustrate particular structures or concepts. The better illustrations are informed by general principles of good visual representation and graphic design, but the linkage between scientific concepts and visual representation is often not deeply principled or algorithmic. Molecules visualisation and databases greatly influence the learning process.

Students were asked to represent molecules using the requested application in practical test. The assesment shows that project scores with the topic of using databases and visualization, a 'good' score profile of the results the two years. The student has good score on Zinc, COD database and avogadro visualisation. These scores have not changed much for two years in a row. Most of the students gave precise answers regarding the definition of molecular structures and database. There are fundamental differences between the search results pages of the three databases, primarily related to the visualization displayed. The interface on the search page and results look more straightforward and intuitive, showing that mastery of the application is followed by knowledge of the database. Databases for small molecules, such as those of ZINC20 and COD, do not use the complex interfaces of VMD. We found that asking students to use a ZINC20 or COD database and then displaying the crystal structure required fewer additional instructions from written instructions as project implementation instructions compared to what was done with PDB. Students mastered the Avogadro application more than VMD. Students generally understand how to get data (computer files) from a database and then the visual constructions requested by the questions.

Acknowledgments: DIPA Badan Layanan Umum Universitas Sriwijaya 2020 funded the writing of this article with SP DIPA-023.17.2.677515 / 2020, revision 01, date March 16 2020. In accordance with the Rector's Decree No: 0685 / UN9 / SK. BUK.KP/2020/ on date July 15 2020.

Conflicts of Interest: The authors declare no conflict of interest..

REFERENCES

Argüello, J. M., & Dempski, R. E. (2020). Fast, Simple, Student Generated Augmented Reality Approach for Protein Visualization in the Classroom and Home Study. *Journal of Chemical Education*, *97*(8), 2327–2331. https://doi.org/10.1021/acs.jchemed.0c00323

Baskin, I. I., & Zhokhova, N. I. (2019). Continuous molecular fields and the concept of molecular co-fields in structure–activity studies. *Future Medicinal Chemistry*, *11*(20), 2701–2713. https://doi.org/10.4155/fmc-2018-0360

Bird, C. L., & Frey, J. G. (2016). From Big Data to Chemical Information. *Chemistry International*, *38*(3–4). https://doi.org/10.1515/ci-2016-3-424

Burgin, S. R., Oramous, J., Kaminski, M., Stocker, L., & Moradi, M. (2018). High school biology students use of visual molecular dynamics as an authentic tool for learning about modeling as a professional scientific practice: Authentic Tool for Learning About Modeling as a

Professional Scientific Practice. *Biochemistry and Molecular Biology Education*, *46*(3), 230–236. https://doi.org/10.1002/bmb.21113

Edwards, T., Foloppe, N., Harris, S. A., & Wells, G. (2021). The future of biomolecular simulation in the pharmaceutical industry: What we can learn from aerodynamics modelling and weather prediction. Part 1. understanding the physical and computational complexity of *in silico* drug design. *Acta Crystallographica Section D Structural Biology*, *77*(11), 1348–1356. https://doi.org/10.1107/S2059798321009712

Fetaji, B., Fetaji, M., Ebibi, M., & Kera, S. (2018). Analyses of Impacting Factors of ICT in Education Management: Case Study. *International Journal of Modern Education and Computer Science*, *10*(2), 26–34. https://doi.org/10.5815/ijmecs.2018.02.03

Fitrani Dinda Fadhilah, Fitri Handayani Harahap, Nur Zarit Sofia, Suhendri Prayoga, & Muhammad Taufik Ihsan. (2021). The Utilization of Information Technology as Learning Media. *Jurnal Riset Dan Inovasi Pembelajaran*, *1*(2), 164–173. https://doi.org/10.51574/jrip.v1i2.48

Gavezzotti, A., & Presti, L. L. (2016). Building Blocks of Crystal Engineering: A Large-Database Study of the Intermolecular Approach between C–H Donor Groups and O, N, Cl, or F Acceptors in Organic Crystals. *Crystal Growth & Design*, *16*(5), 2952–2962. https://doi.org/10.1021/acs.cgd.6b00305

Gražulis, S., Chateigner, D., Downs, R. T., Yokochi, A. F. T., Quirós, M., Lutterotti, L., Manakova, E., Butkus, J., Moeck, P., & Le Bail, A. (2009). Crystallography Open Database – an open-access collection of crystal structures. *Journal of Applied Crystallography*, *42*(4), 726–729. https://doi.org/10.1107/S0021889809016690

Heller, S. R., McNaught, A., Pletnev, I., Stein, S., & Tchekhovskoi, D. (2015). InChI, the IUPAC International Chemical Identifier. *Journal of Cheminformatics*, *7*(1), 23. https://doi.org/10.1186/s13321-015-0068-4

Higgins, S. G., Nogiwa-Valdez, A. A., & Stevens, M. M. (2022). Considerations for implementing electronic laboratory notebooks in an academic research environment. *Nature Protocols*, *17*(2), 179–189. https://doi.org/10.1038/s41596-021-00645-8

Hossain, S. (2019). *Visualization of Bioinformatics Data with Dash Bio*. 126–133. https://doi.org/10.25080/Majora-7ddc1dd1-012

Irwin, J. J., Tang, K. G., Young, J., Dandarchuluun, C., Wong, B. R., Khurelbaatar, M., Moroz, Y. S., Mayfield, J., & Sayle, R. A. (2020). ZINC20—A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *Journal of Chemical Information and Modeling*, *60*(12), 6065–6073. https://doi.org/10.1021/acs.jcim.0c00675

Karonen, M., Murtonen, M., Södervik, I., Manninen, M., & Salomäki, M. (2021). Heuristics Hindering the Development of Understanding of Molecular Structures in University Level Chemistry Education: The Lewis Structure as an Example. *Education Sciences*, *11*(6), 258. https://doi.org/10.3390/educsci11060258

Kiernan, N. A., Manches, A., & Seery, M. K. (2021). The role of visuospatial thinking in students' predictions of molecular geometry. *Chemistry Education Research and Practice*, *22*(3), 626–639. https://doi.org/10.1039/D0RP00354A

Kozlíková, B., Krone, M., Falk, M., Lindow, N., Baaden, M., Baum, D., Viola, I., Parulek, J., & Hege, H.-C. (2017). Visualization of Biomolecular Structures: State of the Art Revisited: Visualization of Biomolecular Structures. *Computer Graphics Forum*, *36*(8), 178–204. https://doi.org/10.1111/cgf.13072

Llanos, E. J., Leal, W., Luu, D. H., Jost, J., Stadler, P. F., & Restrepo, G. (2019). Exploration of the chemical space and its three historical regimes. *Proceedings of the National Academy of Sciences*, *116*(26), 12660–12665. https://doi.org/10.1073/pnas.1816039116

Popova, M., Bretz, S. L., & Hartley, C. S. (2016). Visualizing Molecular Chirality in the Organic Chemistry Laboratory Using Cholesteric Liquid Crystals. *Journal of Chemical Education*, *93*(6), 1096–1099. https://doi.org/10.1021/acs.jchemed.5b00704

Rayan, B., & Rayan, A. (2017). Avogadro Program for Chemistry Education: To What Extent can Molecular Visualization and Three-dimensional Simulations Enhance Meaningful Chemistry Learning? *World Journal of Chemical Education*, *5*(4), 136–141. https://doi.org/10.12691/wjce-5-4-4

Simmons, C. R., Zhang, F., Birktoft, J. J., Qi, X., Han, D., Liu, Y., Sha, R., Abdallah, H. O., Hernandez, C., Ohayon, Y. P., Seeman, N. C., & Yan, H. (2016). Construction and Structure Determination of a Three-Dimensional DNA Crystal. *Journal of the American Chemical Society*, *138*(31), 10047–10054. https://doi.org/10.1021/jacs.6b06508

Stella, E., Agosti, I., Di Blas, N., Finazzi, M., Lanzi, P. L., & Loiacono, D. (2022). A virtual reality classroom to teach and explore crystal solid state structures. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-022-13410-0

Zardecki, C., Dutta, S., Goodsell, D. S., Lowe, R., Voigt, M., & Burley, S. K. (2022). PDB -101: Educational resources supporting molecular explorations through biology and medicine. *Protein Science*, *31*(1), 129–140. https://doi.org/10.1002/pro.4200