**Adult and Higher Education (2024)**
**Clausius Scientific Press, Canada**

**DOI: 10.23977/aduhe.2024.060717**
**ISSN 2523-5826 Vol. 6 Num. 7**

# *CDIO in Bioinformatics Curriculum: A Case Study on Sequence Alignment*

**Jinming Gao[a], Sijia Zhang[b], Dongsheng Chen[c], Xianzhao Kan[d,*]**

*College of Life Sciences, Anhui Normal University, Wuhu, 241000, China*
*[a]gaojming@ahnu.edu.cn, [b]zhangsijia@ahnu.edu.cn, [c]cds2001@ahnu.edu.cn,*
*[d]xianzhao@ahnu.edu.cn*
*[*]Corresponding author*

***Keywords:*** CDIO, Bioinformatics Curriculum, Sequence alignment

***Abstract:*** Bioinformatics, which combines biology, computer science, mathematics and information technology to understand biological data, is becoming more and more important in modern life science and medical research. However, traditional bioinformatics education often focuses on theoretical knowledge, while ignoring practical skills. The CDIO framework, which stands for Conceive, Design, Implement, and Operate, is an innovative educational method to cultivate the next generation of engineers. In this study, we investigate how the CDIO engineering education framework can improve the bioinformatics curriculum, with a focus on teaching sequence alignment. We designed four teaching modules. These modules were created based on the CDIO framework, emphasizing both theoretical knowledge and practical skills. Through these CDIO-based methods, we ensure that students can gain essential skills for careers in bioinformatics and related areas.

## 1. Introduction

The field of bioinformatics, which combines biology, computer science, mathematics, and information technology to understand biological data [1], is becoming more and more important in modern life science and medical research. Over time, bioinformatics education has undergone dramatic changes, moving from simple workshops to formal academic qualifications [2]. Indeed, traditional bioinformatics education often focuses on theoretical knowledge, while ignoring practical skills. Currently, gene sequencing technologies are developing rapidly, from first-generation sequencing methods to high-throughput next-generation sequencing, resulting in big biological data. It is well-known that sequence alignment plays a crucial role in examining these large biological datasets, enabling biologists to identify similarities, differences, and evolutionary relationships among DNA, RNA, and protein sequences. That is to say, sequence alignment is an important component of bioinformatics courses.

The CDIO framework, which stands for Conceive, Design, Implement, and Operate, was originally developed by the Massachusetts Institute of Technology (MIT). It is an innovative educational method to cultivate the next generation of engineers. Across the world, more than 200 institutions have adopted CDIO as the framework for their curriculum development (www.cdio.org/cdio-organization). Introducing CDIO into bioinformatics education can help students better understand core concepts and enhance their problem-solving skills for real-world challenges. In this study, we investigate how the CDIO engineering education framework can improve the bioinformatics curriculum, with a focus on teaching sequence alignment.

## 2. Challenges in Bioinformatics Education

Around the mid-1980s, the term “bioinformatics” was introduced to describe the use of information technology in biological research. An appropriate definition of bioinformatics is “the science of how information is generated, transmitted, received, stored, processed, and interpreted in biological systems” [3]. With bioinformatics becoming vital in life sciences, it is crucial to address the challenges in teaching life scientists the skills they need. Bioinformatics education refers to the study and teaching of bioinformatics concepts, techniques, and applications. The study of bioinformatics includes elements from biology, computer science, mathematics, and statistics.

Isik et al. [4] presented several main challenges in bioinformatics education: fundamental knowledge and skills, lifelong learning, educator development, technology advancement, ethical and legal considerations, inclusivity and equity, and sustainability. For example, effective bioinformatics education and training require highly skilled teachers. However, because bioinformatics is a young and multidisciplinary field, many teachers may lack formal bioinformatics training. Thus, educator development is one of the grand challenges. Furthermore, another primary challenge is “Ethics in Bioinformatics” [5]. There are many ethical issues in the field of bioinformatics, such as individual privacy, data ownership, and the access and use of data. However, current bioinformatics education often neglects ethical considerations. Therefore, there is an urgent need for curriculum reform to incorporate ethical education.

## 3. Progress in Sequence Alignment

In bioinformatics, sequence alignment tools are vital for analyzing and comparing biological sequences, including DNA, RNA, and proteins. These tools can identify similarities, analyze phylogenetic relationships, and infer functional and structural characteristics of these molecules. With the rapid increase in biological data, efficient and accurate alignment tools are becoming more important. Sequence alignment involves comparing pairwise or multiple nucleotide or protein sequences against a reference. Based on the scope of the comparison, pairwise alignment can be further divided into global alignment and local alignment. Global alignment compares entire sequences, while local alignment focuses on finding similar regions within them. Pairwise alignment algorithms, such as Needleman-Wunsch for global alignment, were introduced by Needleman and Wunsch in 1970 [6]. In contrast, the Smith-Waterman algorithm is used for local alignment [7].

Some popular pairwise sequence alignment tools can be found in the JOB Dispatcher of EMBL’s European Bioinformatics Institute [8], including: EMBOSS Needle, EMBOSS Stretcher, GGEARCH2SEQ for global alignment, and EMBOSS Water, EMBOSS Matcher, LALIGN, SSEARCH2SEQ for local alignment. Besides, FASTA [9] and NCBI BLAST [10] are popular tools that employ heuristic algorithms. For multiple sequence alignment, tools such as Clustal Omega, EMBOSS Cons, Kalign, MAFFT, MUSCLE, Mview, T-Coffee, WebPPANK are commonly used. In bioinformatics education, teaching students how to select appropriate sequence alignment methods is crucial. Students need to understand the strengths and limitations of different alignment tools.

## 4. CDIO Framework and Its Application

The CDIO Framework is an innovative educational model specifically for engineering education. It was developed at MIT in collaboration with Chalmers University of Technology, KTH Royal Institute of Technology, and Linköping University. The CDIO model consists of four components: Conceive (defining needs), Design (creating detailed plans), Implement (building and testing the solution), and Operate (deploying and maintaining the product) [11]. The CDIO Standards encompass seven areas: mission, curriculum, practical work, pedagogy, staff training, goals, and assessment [11]. Furthermore, the CDIO syllabus provides a framework for engineering education: personal and professional skills, interpersonal skills, and CDIO abilities. Moreover, the CDIO Initiative Mission is “Building community capacity to make an open flexible and evolving framework for the advancement of engineering education by an inclusive collaboration and a sharing of effective practices for local impact” (cdio.org/cdio-organization).

Currently, the CDIO Framework has been applied across various engineering disciplines, including architectural education [12], as well as in the development of student skills and attributes [13], faculty development programs [14], and the evaluation of Swedish engineering degree programmes [15]. Additionally, it has been used in project-based learning in higher education [16], technological innovation [17], and specific engineering courses such as programming [18], mechanical engineering [19], chemical engineering [20], and capstone project [21]. Furthermore, it has been employed to address broader topics such as environmental sustainability [22], professional skills [23], and the master’s training [24]. The CDIO Framework provides a practical, structured approach to engineering education, with principles that are highly applicable to bioinformatics, particularly in teaching sequence alignment.

## 5. Case Study: Applying the CDIO Framework to Sequence Alignment

### 5.1. Description of the Bioinformatics Curriculum in Anhui Normal University

The College of Life Sciences at Anhui Normal University, China, was founded in 1930. The Institute of Bioinformatics was established in 2008, with faculty from the College of Life Sciences and the College of Computer Sciences. That same year, a bioinformatics course was introduced. It was optional at the beginning, but is now mandatory. Over 300 students take this course every year. The content of the bioinformatics curriculum at our university includes the history of bioinformatics, biological databases, sequence alignment, protein structure prediction, genomics, proteomics, phylogenetics (as detailed in Table 1). Currently, our teaching team has six members, all of whom hold PhDs, and graduated from various institutes of the Chinese Academy of Sciences and universities both in China and abroad. This course is available to seventh-semester undergraduates in Biological Sciences, Biotechnology, and Biopharmaceuticals.

### 5.2. Conceiving Learning Objectives Related to Sequence Alignment

By comparing the similarity between two or more sequences, we can determine if they share homology. When aligning multiple protein or nucleic acid sequences, we can identify conserved motifs that share common biological functions. These conserved motifs can also provide insights into the functions of novel sequences. The main objectives of sequence alignment learning are detailed in Table 2.

Table 1: The main content of bioinformatics curriculum in Anhui Normal University.

| Chapter | Title | Main content |
|---|---|---|
| Chapter 1 | The concept of bioinformatics and its development history | The development history, research areas, major applications, and challenges of bioinformatics |
| Chapter 2 | Biological databases and their retrieval | The concepts of primary databases and secondary databases, the content and structure of biological databases, and the search of biological databases |
| Chapter 3 | The principle of sequence alignment | The concepts related to sequence alignment, scoring methods, and alignment algorithms, as well as tools for pairwise and multiple sequence alignment |
| Chapter 4 | Protein structure analysis | The levels of protein structure organization, the determination and theoretical prediction of protein structures, and protein folding and disease. |
| Chapter 5 | Annotation of eukaryotic genomes | Annotation of protein-coding genes, RNA genes, repetitive genes, pseudogenes, and case-based teaching |
| Chapter 6 | Proteomics | Overview of proteomics, large-scale protein separation and identification techniques, protein post-translational modifications, protein sorting, and protein interactions. |
| Chapter 7 | Molecular evolution and phylogenetics | Molecular evolution and phylogenetics, methods for constructing molecular phylogenetic trees, and the construction and application of phylogenetic trees |
| Chapter 8 | Foundations of bioinformatics programming | Linux operating system, programming languages in bioinformatics, SQL and database programming, and parallel computing |

Table 2: Sequence alignment learning objectives.

| Main Objectives | Detailed Explanation |
|---|---|
| Understand the principle of sequence alignment | The principle of sequence alignment is to compare two or more biological sequences (DNA, RNA, or protein) to identify similarities and differences. |
| Learn various types of sequence alignment | Understand the difference between global alignment and local alignment, and their applications. |
| Understand the primary algorithms for sequence alignment | Learn the basic principles and applications of dynamic programming algorithms (e.g., Needleman-Wunsch, Smith-Waterman) and heuristic algorithms (e.g., BLAST). |
| Grasp the concepts of scoring matrices and gap penalties | Grasp the selection of different scoring matrices (e.g., PAM, BLOSUM) in different biological problems, and the role of gap penalties in sequence alignment. |
| Learn how to use sequence alignment tools | Be proficient in using common sequence alignment software (e.g., BLAST, ClustalW) |
| Know how to interpret sequence alignment results | Understand the biological significance of similarities, differences, conserved regions in results. |

### 5.3. Design of CDIO-Based Course Modules for Sequence Alignment

To help students at Anhui Normal University learn about sequence alignment, we have designed four teaching modules. These modules were created based on the CDIO framework, emphasizing both theoretical knowledge and practical skills. The modules include:

(1) Module 1: Conceive - sequence alignment fundamentals;
(2) Module 2: Design - alignment algorithms;
(3) Module 3: Implement - coding and testing alignment tools;
(4) Module 4: Operate - deploying and maintaining alignment tools.

Furthermore, we also present the objectives and activities for each module (see Figure 1).

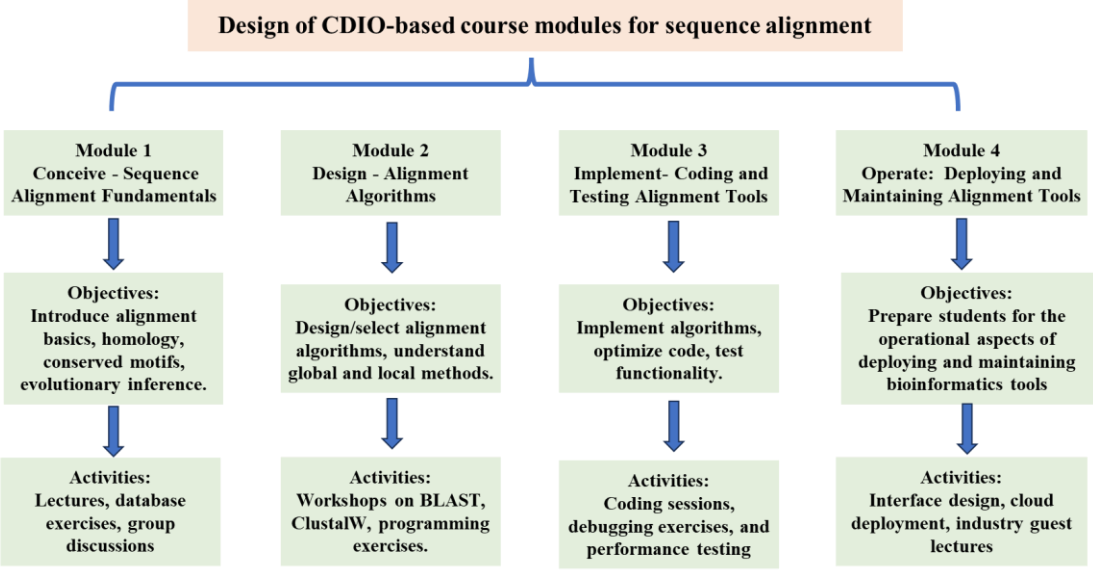


Figure 1: Design of CDIO-based course modules for sequence alignment

### 5.4. Implementing Teaching Strategies Using Sequence Alignment Tools

Our teaching strategies focus on active learning and problem-solving skills. For example, we employ a flipped classroom model, where students read the lesson materials before class, and then work with others to practice using sequence tools during class.

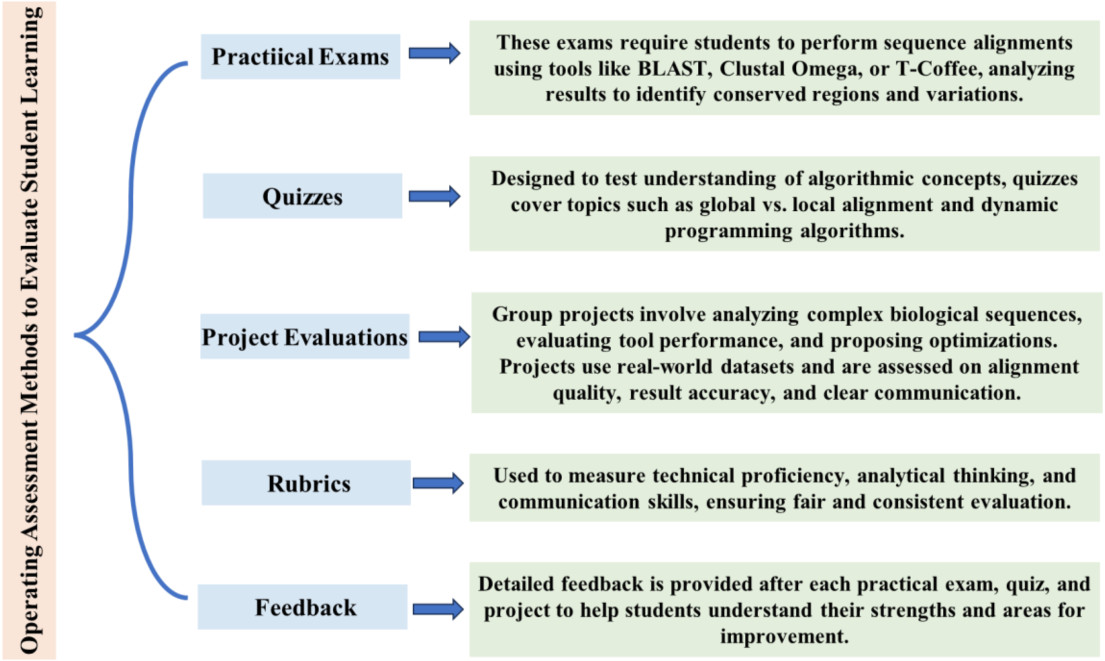
Before class on teaching sequence alignment tools, students should learn: (1) the website and functions of FASTA tools, (2) the online and local versions of BLAST, (3) the usage of Clustal Omega, T-Coffee, MultAlin, and MAFFT software.

During class, we provide several questions as follows: (1) Which software is suitable for pairwise sequence alignment, and which is suitable for multiple sequence alignment? (2) Which software is suitable for global alignment, and which is suitable for local alignment? (3) Between FASTA tools and BLAST tools, which one computes faster? (4) What are the differences among blastn, blastp, tblastx, tblastn, and tblastx? (5) What advanced BLAST tools are available? (6) What are the characteristics and differences among various multiple sequence alignment tools? In the classroom, students will address these questions through hands-on practice, group discussions, and whole-class discussions. We then provide summary comments and feedback.

### 5.5. Operating Assessment Methods to Evaluate Student Learning

To better evaluate students' learning in sequence alignment, we employed assessment methods, including practical exams, quizzes, etc. The detailed evaluation system can be seen in Figure 2.

These assessments check both students' practical skills and their ability to understand and explain results. By combining these assessment methods, we ensure that students studying Biological Sciences, Biotechnology, and Biopharmaceuticals gain essential skills for careers in bioinformatics and related areas.



These exams require students to perform sequence alignments using tools like BLAST, Clustal Omega, or T-Coffee, analyzing results to identify conserved regions and variations.

Designed to test understanding of algorithmic concepts, quizzes cover topics such as global vs. local alignment and dynamic programming algorithms.

Group projects involve analyzing complex biological sequences, evaluating tool performance, and proposing optimizations. Projects use real-world datasets and are assessed on alignment quality, result accuracy, and clear communication.

Used to measure technical proficiency, analytical thinking, and communication skills, ensuring fair and consistent evaluation.

Detailed feedback is provided after each practical exam, quiz, and project to help students understand their strengths and areas for improvement.

Figure 2: Assessment strategies for sequence alignment learning

## 6. Conclusions

Over time, bioinformatics education has undergone dramatic changes, and faces challenges such as practical skills and ethical considerations. This study demonstrates that the CDIO framework can improve bioinformatics education, especially for teaching sequence alignment. By integrating the CDIO steps of Conceiving, Designing, Implementing, and Operating into the curriculum, we provided a comprehensive learning experience for students at the College of Life Sciences, Anhui Normal University. We designed four teaching modules. These modules were created based on the CDIO framework, emphasizing both theoretical knowledge and practical skills. Moreover, we employed a flipped classroom model, where students read the lesson materials before class, and then work with others to practice using sequence alignment tools during class. To better evaluate students' learning in sequence alignment, we employed assessment methods, including practical exams, quizzes, etc. Through CDIO-based methods, we ensure that students can gain essential skills for careers in bioinformatics and related areas.

## Acknowledgements

This work was financially supported by Anhui Provincial Quality Engineering Project for Higher Education Institutions (2021xnfzxm040, 2022jyxm538, 2023jyxm0175 and 2023qyw/sysfkc010).

## References

*[1] P. K. Akalın (2006) Introduction to bioinformatics. Molecular nutrition & food research. 50: 610-619.*
*[2] S. Ranganathan (2005) Bioinformatics education—perspectives and challenges. PLoS computational biology. 1: e52.*

*[3] J. Ramsden (2023) Bioinformatics: an introduction. Springer Nature.*

*[4] E. B. Işık, M. D. Brazas, R. Schwartz, B. Gaeta, P. M. Palagi, C. W. van Gelder, P. Suravajhala, H. Singh, S. L. Morgan and H. Zahroh (2023) Grand challenges in bioinformatics education and training. Nature Biotechnology. 41: 1171-1174.*

*[5] B. Taneri (2011) Is there Room for Ethics within Bioinformatics Education? J. Comput. Biol. 18: 907-916.*

*[6] S. B. Needleman and C. D. Wunsch (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of molecular biology. 48: 443-453.*

*[7] T. F. Smith and M. S. Waterman (1981) Identification of common molecular subsequences. Journal of molecular biology. 147: 195-197.*

*[8] F. Madeira, N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A. R. Tivey, R. Lopez and S. Butcher (2024) The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. Nucleic Acids Research. gkae241.*

*[9] W. R. Pearson (1999) Flexible sequence similarity searching with the FASTA3 program package. Bioinformatics methods and protocols. 185-219.*

*[10] S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990) Basic local alignment search tool. Journal of molecular biology. 215: 403-410.*

*[11] E. F. Crawley, J. Malmqvist, S. Östlund, D. R. Brodeur, K. Edström, E. F. Crawley, J. Malmqvist, S. Östlund, D. R. Brodeur and K. Edström (2014) The CDIO approach. Rethinking Engineering Education: The CDIO Approach. 11-45.*

*[12] L. Nyka, J. Cudzik and K. Szakajło (2020) The CDIO model in architectural education and research by design. World Transactions on Engineering and Technology Education. 18: 85-90.*

*[13] P. Armstrong and R. Niewoehner (2008) The CDIO approach to the development of student skills and attributes. Proceedings of the 4th International CDIO Conference, Hogeschool Gent, Belgium.*

*[14] A. Chuchalin, M. Tayurskaya and J. Malmqvist (2015) Faculty development programme based on CDIO framework. Proceedings of the 2015 International Conference on Interactive Collaborative Learning (ICL).*

*[15] J. Malmqvist, K. Edstrom, S. Gunnarsson and S. Ostlund (2006) The application of CDIO Standards in the evaluation of Swedish engineering degree programmes. World Transactions on Engineering and Technology Education. 5: 361.*

*[16] M. Säisä, S. Määttä and J. Roslöf (2017) Integration of CDIO skills into project-based learning in higher education. Proceedings of the Proceedings of the 13th International CDIO Conference.*

*[17] F. Smulders, A. Kamp and C. Fortin (2018) The CDIO framework and new perspectives on technological innovation. Proceedings of the 14th International CDIO Conference.*

*[18] C.-F. Lai, H.-X. Zhong and P.-S. Chiu (2021) Investigating the impact of a flipped programming course using the DT-CDIO approach. Computers & Education. 173: 104287.*

*[19] A. S. MUSHTAK AL-ATABI and M. AL-OBAIDI (2011) CDIO curriculum for mechanical engineering undergraduate course. Journal of engineering science and technology. 6: 251-259.*

*[20] S.-M. Cheah (2009) Using CDIO to revamp the chemical engineering curriculum. Proceedings of the 5th International CDIO Conference.*

*[21] P. Armstrong, R. Kee, R. Kenny and G. Cunningham (2005) A CDIO approach to the final year capstone project. Proceedings of the 1st Annual CDIO Conference.*

*[22] M. K. Wedel, J. Malmqvist, M. Arehag and M. Svanström (2008) Implementing engineering education for environmental sustainability into CDIO programs. Proceedings of the Proceeding of the 4th International CDIO Conference.*

*[23] L. J. Leslie, P. C. Gorman and S. Junaid (2021) Conceive-design-implement-operate (cdio) as an effective learning framework for embedding professional skills. International Journal of Engineering Education. 37: 1289-1299.*

*[24] J. Antokhina, V. Olenev and Y. Sheynin (2016) CDIO implementation experience for the masters training at SUAI. Proceedings of the Proceedings of the 12th International CDIO Conference, Turku University of Applied Sciences, Turku, Finland.*