



TH Bingen Technical University of Applied Sciences
Department 2 - Technology, Informatics and Economics
Applied Bioinformatics (B.Sc.)

Titel

Hausarbeit Wissenschaftliches Arbeiten
Abgegeben am: ???
Name: ???
Matrikelnummer: ???

The Abstract is a short summary of the scientific publication that usually contains a short description of the publications topic to help the reader to decide whether to read the full paper or not. Optimally the reader is convinced to read the full paper after the author generated enough interest. The Abstract prepares the readers for the detailed an in-depth processes of the topic. At last, an abstract helps the reader to remember key points from the publication. Citations must be done in this style:

Abbildungsverzeichnis

1	Example of a captioned figure	3
2	An example of a subdivided figure.	5
3	An example of an unsubdivided figure.	6

Tabellenverzeichnis

1	Distribution of values tested during simulated annealing in 4th quartile of high scoring parameter sets	4
---	--	---

Inhaltsverzeichnis

1	Introduction	1
1.1	questions to answer	1
1.2	Introduction style	1
1.2.1	This is a sub-section	1
2	Material	2
2.1	List of topics	2
3	Methods	3
3.1	Reproduction	3
3.2	Formula examples	3
4	Results	5
4.1	Envelope Variants	5
4.2	Code Example	6
4.3	Example of using a reference to another section	7
5	Discussion	8

1 Introduction

1.1 questions to answer

The Introduction mostly answers the following questions in a short manner:

- What is the topic of this scientific publication?
- What is the current point of research?
- What materials and which methods were used to work on the topic?
- What is the hypothesis, that was to confirm or deny?
- This is an example for a citation ([1]).

1.2 Introduction style

An Introduction is written in plain text, and is usually written as one paragraph of continuing text, although it is also possible to subdivide the text into different partitions.

1.2.1 This is a sub-section

Lorem ipsum bla foo bar...

2 Material

2.1 List of topics

Multilevel List of the Data and Substances content:

1. What kind of data has been used and in which version?

- Database XY

XY's function and description: the database was used in the version 2002002.2.456 and only contains the latest premium curated data, just like SILVA f.e.

Genomes: the used genomes were the genes C137-A, C137-B and C137-C

2. Which substances and organisms have been used?

- Phage in genetically modified version

Function: the phages function was to carry a plasmid ring with modified rna into the victim organism

Variants: Version 1442.312.1.666

3 Methods

3.1 Reproduction

Throughout the whole paper there is a need to describe every detail and every step in a precise manner to make sure all created results can be reproduced by others. First and foremost this reproduction demand has to be ensured in the methods chapter.

Also very important are:

- A chronological order of all steps (including software versions, details, data-handling, ..)
- What method was used to solve each part of the task?
- What data versions created which results?

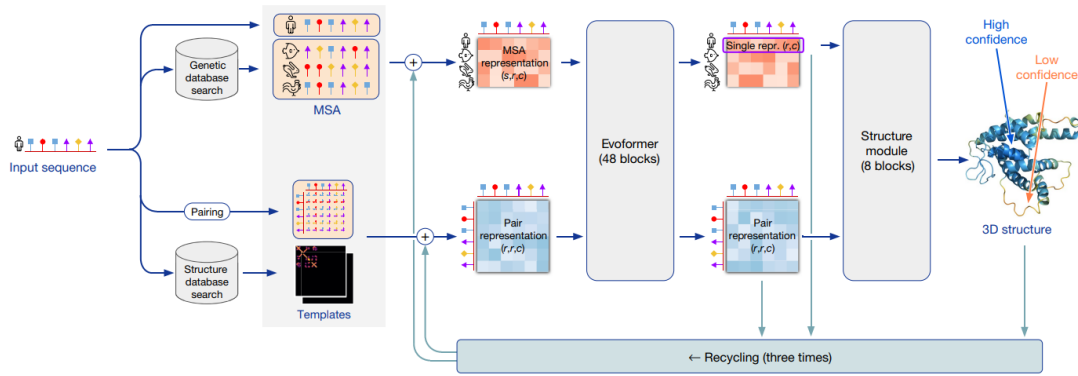


Abbildung 1: Example of a captiones figure

This is a single-line code example:

```
$ docker run --rm --gpus all nvidia/cuda:11.0-base nvidia-smi
```

3.2 Formula examples

Scores below 0.2 belong to randomly selected unrelated proteins, according to stringent statistics of structures in the PDB, while those with more than 0.5 generally have the same protein fold ([1]) oder [1]. The TM-score is calculated defined by:

$$TM - score = \max \left[\frac{1}{L_{target}} \sum_i^{L_{common}} \frac{1}{1 + (\frac{d_i}{d_0(L_{target})})^2} \right]$$

Where L_{target} is the length of the amino acid sequence of the target protein, and L_{common} is the number of residues that appear in both the template and target structures, d_i is the distance between the i th pair of residues in the template and target structures, and $d_0(L_{target})$ is a distance scale that normalizes distances and is defined by:

$$d_0(L_{target}) = 1.24\sqrt[3]{L_{target} - 15 - 1.8}$$

This is a reference to the labeled (numbered) formula [3.2] When comparing two protein structures that have the same residue order, L_{common} reads from the C-alpha order number of the structure files. The RMSD value is defined by:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta i^2}$$

Where δi is the distance between atom i and either a reference structure or the mean position of the N equivalent atoms.

Tabelle 1: Distribution of values tested during simulated annealing in 4th quartile of high scoring parameter sets (mean F2-Scores > 0.6454).

Parameter	Minimum	1st Quartile	Median	Mean
α	0.1000	0.2000	0.4000	0.4692
β	0.0476	0.3704	0.4545	0.4622
ω	0.0476	0.1250	0.2105	0.2324
σ	0.0476	0.1875	0.3077	0.3054
Sprot- w	10.0000	20.0000	30.0000	41.02009738
Sprot- δ	0.1000	0.3000	0.5000	0.5365
trEMBL- w	10.0000	50.0000	70.0000	67.10005704
trEMBL- δ	0.1000	0.5000	0.7000	0.6823
TAIR- w	10.0000	30.0000	50.0000	53.79005853
TAIR- δ	0.1000	0.3000	0.5000	0.5392

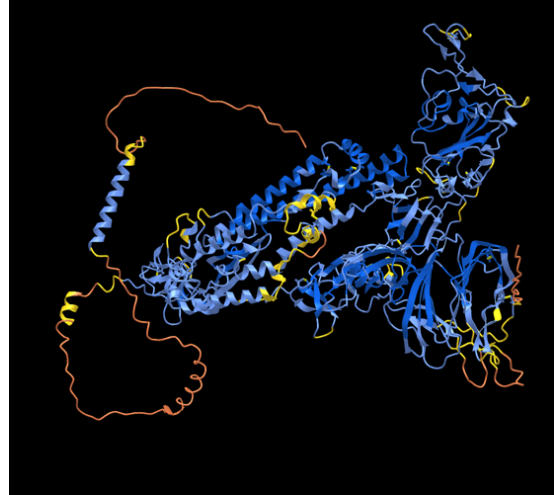
This is an example on how to reference a table that has the `\label{tbl:value_table}` by using `\ref{tbl:value_table}`: The estimated values are listed in table 1.

4 Results

4.1 Envelope Variants



(a) Left part of a figure partitioned into subfigures.



(b) Right part of a figure partitioned into subfigures.

Abbildung 2: An example of a subdivided figure.

As we have labeled the above figure containing subfigures, we can reference it using 2.

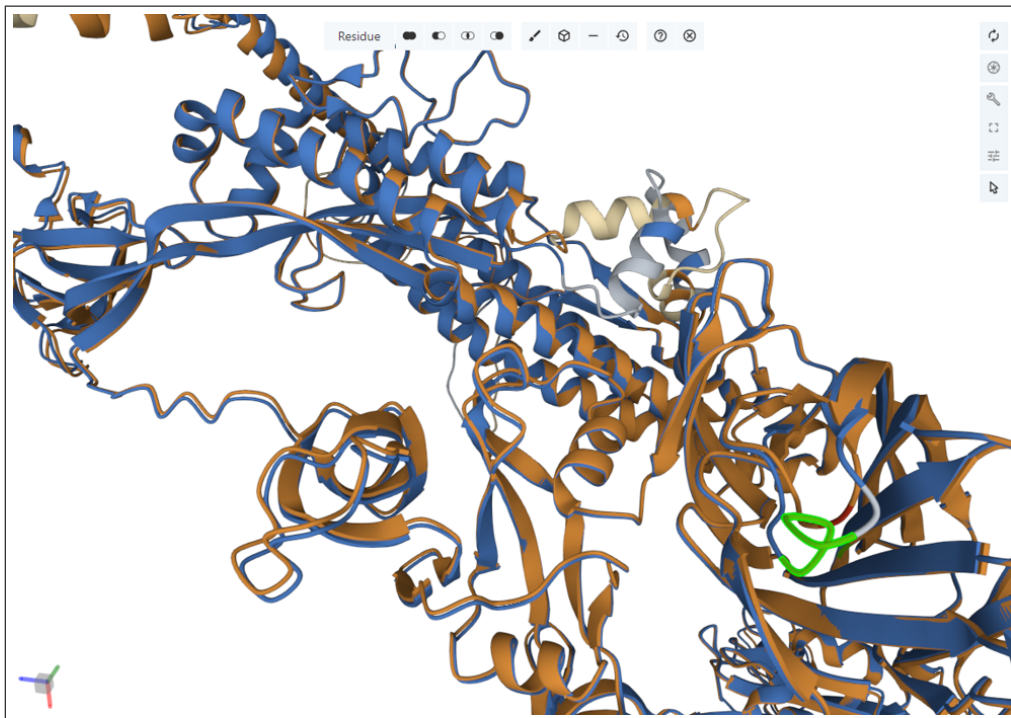


Abbildung 3: An example of an unsubdivided figure.

Similarly, we can reference the unsubdivided figure using 3.

4.2 Code Example

The results of the study include the generated data, outcomes of statistical tests, and visual evaluations such as plant leaf imagery. This section focuses on presenting these results objectively without interpretation. Incremental values can be depicted using tables and figures where necessary.

Below is an example of how to insert an algorithm in LaTeX:

Algorithm 1: Mutation Probability Lookup Algorithm

Data:

- g : The GO term to compute the mutation probability lookup table for.
- \mathcal{H} : The *ordered* set of homologous protein pairs p_i , where
 - at least one member has annotation g ,
 - with \mathcal{G}_x as the set of GO term annotations of protein x :
 - $p_i = \{p_i^1, p_i^2 \mid \exists k \in \{1, 2\} : g \in \mathcal{G}_{p_i^k}\}$, and
 - each pair p_i has a sequence distance d_i such that: $d_{i-1} \leq d_i$.

Result:

- M : Ordered set of mutation probabilities $P(g^{mut} \mid d_k)$, with
 - $m_{k-1} \leq m_k$ for $d_{k-1} \leq d_k$.

```
1 Initialize:
2  $n_{not\_sharing} \leftarrow 0$ 
3  $n_{all} \leftarrow 0$ 
4  $m_{candidate} \leftarrow 0$ 
5  $m_{current} \leftarrow 0$ 
6 foreach  $p_i \in \mathcal{H}$ ,  $1 \leq i \leq |\mathcal{H}|$  do
7   Increment  $n_{all}$  by 1.
8   if  $\exists k \in \{1, 2\} : g \notin \mathcal{G}_{p_i^k}$  then
9     Increment  $n_{not\_sharing}$  by 1.
10   $m_{candidate} \leftarrow \frac{n_{not\_sharing}}{n_{all}}$ 
11  if  $m_{candidate} > m_{current}$  then
12     $m_{current} \leftarrow m_{candidate}$ 
13    Append  $m_{current}$  to  $M$ .
```

4.3 Example of using a reference to another section

This is a reference example. To reference the code example above, simply use `\ref{alg:mutation_lookup}` 1.

Find more information about the materials in section 2.1

(`\ref{Topics_of_Material}`) on page 2 (`\pageref{Topics_of_Material}`).

5 Discussion

- What is the meaning of the results? (interpretation included here)
- How robust is the experiment and what are the weak links?
- How reliable are the results?
- What are possible next steps? \Rightarrow New hypotheses / perspectives / follow-up experiments

Literatur

- [1] John Marston. *Beispieltitel*. 1650. URL: www.beispiel_URL.de (besucht am 27.02.2023).

Erklärung zur Originalität der Arbeit

Hiermit bestätige ich, dass die abgegebene Arbeit das Original ist und von mir ohne weitere Hilfe geschrieben wurde. Wenn Arbeit anderer referenziert oder genutzt wurde, wurde dies angemessen kenntlich gemacht. Meine Arbeit wurde noch nicht bewertet oder veröffentlicht. Die elektronisch abgegebene Version stimmt mit der elektronischen überein.

Unterschrift

Ort und Datum

Erklärung zum Eigentum und Urheberrecht

Hiermit erkläre ich meine Zustimmung, dass die Technische Hochschule Bingen diese Arbeit anderen Studierenden und interessierten Dritten zur Verfügung stellen und in meinem Namen (Robin Ender) veröffentlichen darf.

Unterschrift

Ort und Datum