DEVELOPMENT OF A SOFTWARE APPLICATION FOR HANDLING PROTEIN CRYSTALLISATION CONDITIONS AND TRIALS

D. Markopoulos[1,2], E. S. Manolakos[2] and E.D. Chrysina[1]

[1]National Hellenic Research Foundation, Institute of Organic & Pharmaceutical Chemistry, 48 Vassileos Constantinou Ave., 11635 Athens, Greece; [2] National and Kapodistrian University of Athens, Postgraduate Program "information Technologies in Medicine and Biology", Department of Informatics & Telecommunications, Panepistimiopolis, Ilissia, Athens 15784, Greece.

e-mail: dionisis_m@yahoo.com

Introduction

The new automated high throughput approaches applied in the field of macromolecular crystallization during the last decade have resulted in the generation of a huge amount of information[1]. To meet the emerging need for handling this information a number of tools has been developed. Our work focuses on the organization of commercial crystallization conditions, as well as those derived from a crystallization lab routine, in a database to support the crystal optimization process. We have developed a computer software application for handling protein crystallization conditions and trials. The new tool facilitates organization and storage of crystallization related data in a relational database.

System components & requirements

The application developed comprises a relational database and a database management system (DBMS) to meet user requirements as described in Table 1. The data are imported through user friendly forms and are organized in a repository. Our DBMS provides users the option to submit queries and retrieve data from the database that will be exploited for further crystallization trials. The structure of the system built and the relevant UML model of the database are shown in Figures 1 and 2, respectively.

Table 1

System components	A relational database
	A database management application
User Requirements	Recording & organization of protein crystallization data
	Data mining for optimization of crystallization conditions

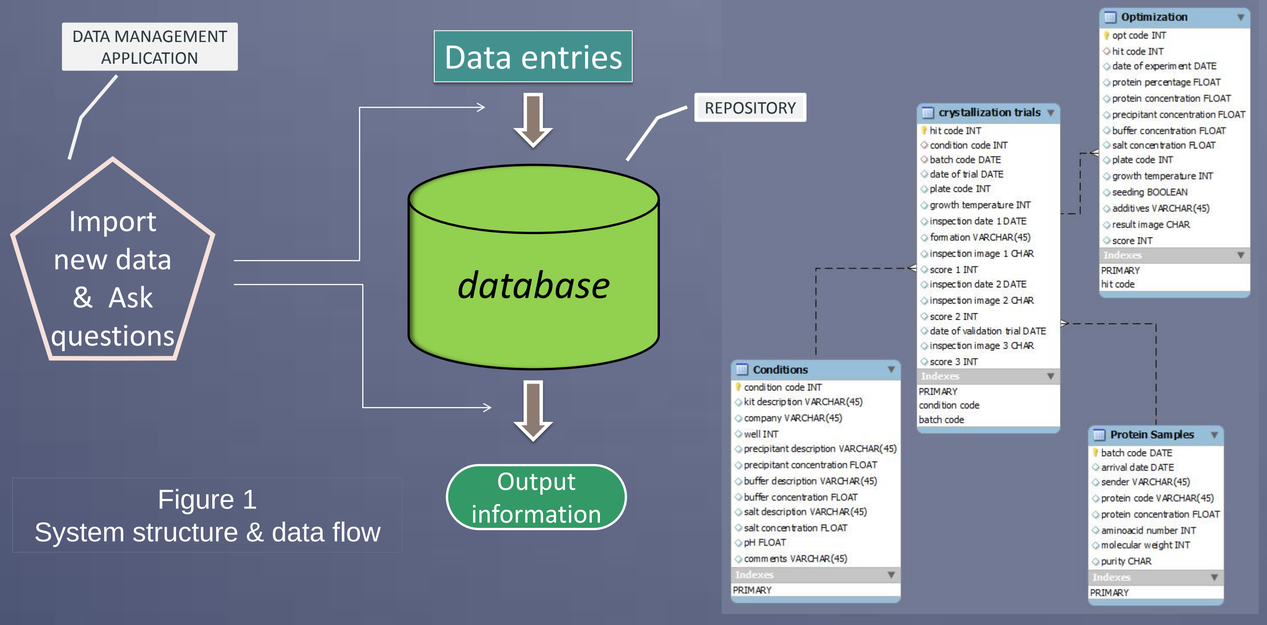


Figure 1
System structure & data flow

Figure 2:
UML model of the database

Front-end environment

The database management application employs a web-based environment as front end, therefore users can run the application on any web browser. Addition of new data through the individual forms and submission of search queries in the database can be very simple tasks. Previews of the forms used for different functions are presented below (Figure 3).

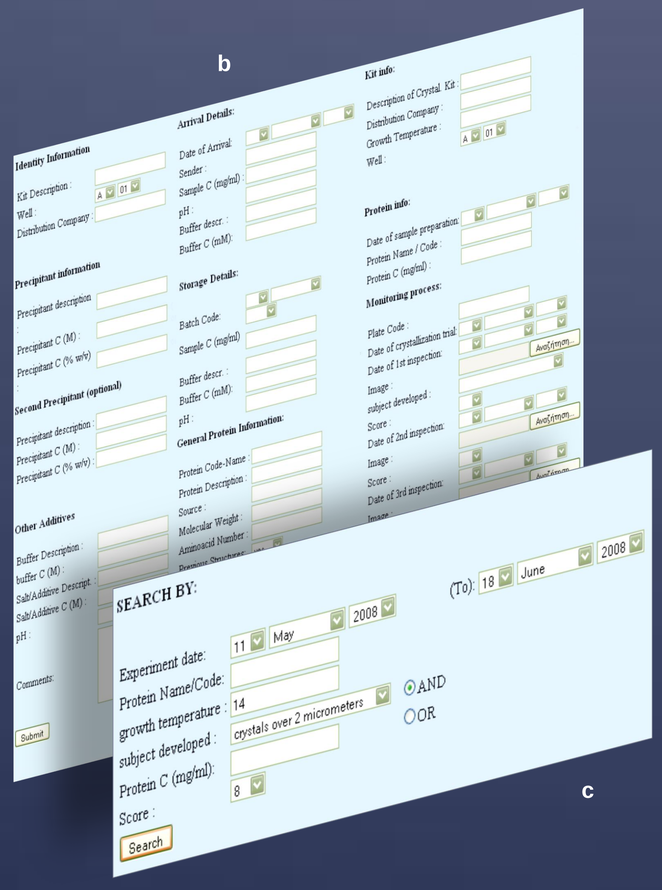
a

Submit New Conditions*** Search for Conditions*** Make corrections*** Go Back

Crystallization Conditions meeting the requirements

Kit Description	Well	Company	Precipitant Description	Precipitant C (mM)	Buffer Description	Buffer C (mM)	Salt/Additive Description (mM)	Salt/Additive C (mM)	pH
Structure Screen 1 & 2 HT-96	A03	jena bioscience	MPD	2.1	sodium sulfate	2	ammonium sulfate	3.1	6.5
Structure Screen 1 & 2 HT-96	A06	jena bioscience	MPD	2.1	Na acetate	2	ammonium acetate	2	4.3
Structure Screen 1 & 2 HT-96	C10	jena bioscience	PEG 2000	2	Na citrate	1	non existent	0	4
Structure Screen 1 & 2 HT-96	A01	molecular dimensions	PEG 1000	0	Na citrate	25	calcium chloride	100	5.4
Structure Screen 1 & 2 HT-96	A02	molecular dimensions	MPD	1.45	nero	3	cinamn	3.4	8

b



c

Figure 3
Screenshots of the front end environment
a. Results of a search query.
b. Submission forms
c. A search query form

Types of Data

The structure of the database emphasizes into 4 different data categories as outlined in Figure 4:

- Commercially available and home-designed screens of crystallization conditions.
- Protein samples used in crystallization trials.
- Optimization conditions explored to improve the quality of the crystal.
- Evaluation scores of the crystallization trials results

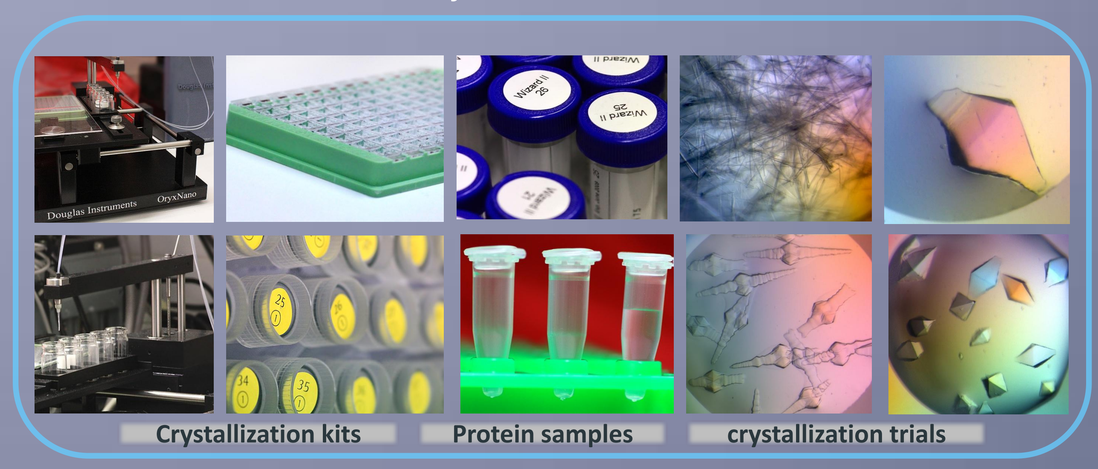


Figure 4
Types of data imported in the database

Search queries

Users can compose their own search queries (Figure 3b), defining a specific range of variables concerning:

- crystallization solution properties (pH, concentration, etc.)
- experimental conditions (temperature, seeding, volume, etc)
- other parameters relevant to a particular experiment (date of crystallization trial / observation, sender, etc.).

Researchers may retrieve valuable information (e.g. crystallization conditions of homologous proteins) in order to design their next set of crystallization experiments.

How to use

Here we describe step by step a convenient method to make use of our application as a tool in the crystallization process (Figure 4).

- Use the available templates to submit your data (protein sample, crystallization conditions, etc.) (step 1).
- Compose your preferred search query (step 2).
  - If you get positive results in your first series of crystallization trials:
    - Use the experiment conditions as a guideline in your query
  - If your crystallization trials have no significant success:
    - Try to form a query based on the properties of your protein.
- Use the results of your search to adjust your next round of crystallization experiments (step 3).
- Do not forget to submit the conditions of your last trial in the database as a new optimization experiment (step 4).

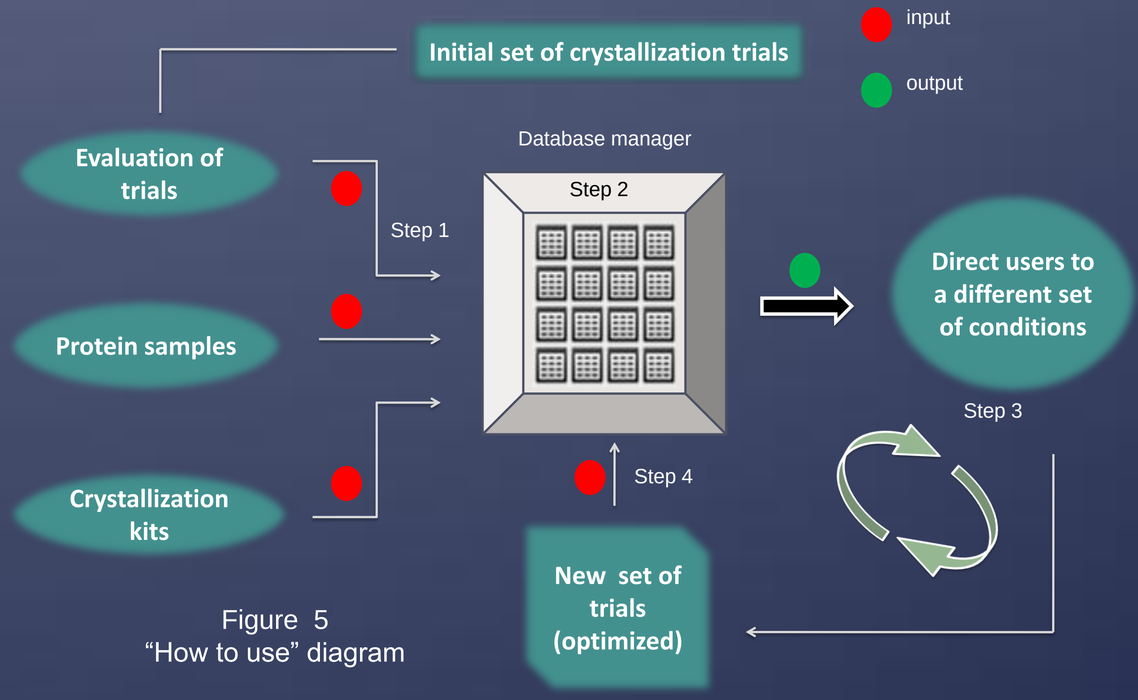


Figure 5
"How to use" diagram

References

1. I.M. Berry et al., Acta Cryst. (2006). D62, 1137–1149
2. C. E. Kundrot, CMLS, Cell. Mol. Life Sci. Vol. 61, 2004

Acknowledgements

This work has been supported by the FP7 Capacities coordination and support actions REGPOT-2008-1-No 230146 'EUROSTRUCT' and REGPOT-2009-1-No 245866 'ARCADE'.