24.4.4. BMCD implementation – web interface

The BMCD is a web-accessible resource available to the crystallographic community through the website at http://wwwbmcd.nist.gov:8080/bmcd/bmcd.html. The current version of the BMCD includes 3547 crystal entries from 2526 biological macromolecules. The web interface provides an easy mechanism for browsing through the data contained in the BMCD. The user can examine the complete list of macromolecule names and tabulations of the number of macromolecules and crystal forms for each source, prosthetic group, space group, chemical addition and crystallization method. In addition, the listing of complete references is available along with a set of general references concerning all aspects of crystal growth.

The web interface offers a number of ways to query the database. For example, the results of precomputed queries are available through the tabulations of chemical additives, space groups, crystallization methods and prosthetic groups mentioned above. The tables provide links to lists of macromolecules and crystal forms that match the database query for these parameters. Allowing the user to enter specific parameter values that must be matched provides another method of searching the database. Examples of this include querying for macromolecules based on their molecular weight. A user may also search for crystal forms of macromolecules that crystallize at a particular temperature, macromolecule concentration and pH. A range of values or a single value of one or all of the parameters may be used to limit the search. For references, queries for specific authors, keywords and journal information are allowed.

24.4.5. Reproducing published crystallization procedures

The BMCD contains the information needed to reproduce the crystallization conditions for a biological macromolecule reported in the literature. This is an activity performed by many laboratories engaged in protein engineering, rational drug design, protein stability and other studies of proteins whose structures have been previously determined. The crystallization of sequence variants, chemically modified derivatives, or ligand–biological macromolecule complexes can sometimes be considered problems that fit into this category. Usually, the reported crystallization conditions of the native macromolecule are the starting points for initiating the crystallization trials. The crystallization of the biological macromolecule may be simple to reproduce, but differences in the isolation and purification procedures, reagents, and crystallization methodology of different laboratories can dramatically influence the results. The crystallization conditions in the database should be considered a good starting point for the search or optimization that will require experiments that vary pH, macromolecule and reagent concentrations, and temperature, along with the crystallization method.

The crystals for one of the isozymes of glutathione S-transferase of rat liver grown from conditions reported in the literature (Sesay *et al.*, 1987) are used to illustrate these points. The original crystallization conditions were for an enzyme isolated from rat liver (entries M0P3 and C13R). However, the enzyme used in the crystallization trials was cloned and expressed in *Escherichia coli*. The crystals of the natural enzyme were grown in 3 to 5 days from vapour-diffusion experiments at 4 °C, with droplets containing a protein concentration of 11.3 mg ml^{-1}, 0.46% β-octylglucoside, 30–37% saturated ammonium sulfate and 0.1 *M* phosphate buffer, pH 6.9 equilibrated against well solutions containing 60–74% ammonium sulfate.

The recombinant enzyme required an optimization of these conditions to produce large single crystals (Ji *et al.*, 1994). The recombinant protein crystallized best at 4.0 °C, with droplets containing a protein concentration of 12 mg ml^{-1}, 0.2% β-octylglucoside, 20–25% saturated ammonium sulfate, 1 m*M* EDTA and 0.025 *M* TrisHCl, pH 8.0 equilibrated against well solutions containing 40–50% ammonium sulfate. Both crystallization protocols required the presence of 1 m*M* (9*R*,10*R*)-9-*S*-glutathionyl-10-hydroxy-9,10-dihydrophenanthrene, a product inhibitor. The recombinant enzyme, β-octylglucoside and ammonium concentrations were adjusted. The pH was varied, with the largest crystals being found at pH 8.0. Thus, TrisHCl was substituted for the phosphate buffer. EDTA was also included as an additive; its

Macromolecule -- subtilisin BPN' : prodomain (M1MT)

Molecule Name : subtilisin BPN' : prodomain
Alias Names :
Genus: Bacillus
Species: amyliloquefaciens
Common Name: recombinant
EC Number : 3.4.21.14
Catalytic Reaction : hydrolysis of proteins and peptide amides
Total Molecular Weight: 36500 **Total No. Subunits:** 2
1. Subunit Molecular Weight: 28000; Subunit Number: 1
2. Subunit Molecular Weight: 8500; Subunit Number: 1

Prosthetic Group :
Remarks : During biosynthesis of subtilisin, the 77 amino acid prodomain is autocatatalytically cleaved.

Total Number of Crystal Entries: 2

1. Crystal C2CJ: A=71.9, B=96.7, C=48.9, alpha=90, beta=90, gamma=90
2. Crystal C2CK: A=74.1, B=77.85, C=57.65, alpha=90, beta=90, gamma=90

Fig. 24.4.3.1. A representative example of a biological macromolecule, subtilisin BPN′: prodomain, entry M1MT in the BMCD.

Crystallization Data for Molecule - subtilisin BPN' : prodomain

Crystal Data (C2CK).
Molecular Name: subtilisin BPN' : prodomain.

Unit cell dimension						
A	B	C	Alpha	beta	gamma	Z
74.1	77.85	57.65	90	90	90	4

Space group:
P2<1>2<1>2 - Orthorhombic
Crystal Density: 0
Diffraction limit (Angstroms): 2
Diffraction life time (Hours): 0
Crystal dimensions in mm: 0.1 x 0.5 x 0.5

Crystallization Data.
Method(s) used for the crystallization: vapor diffusion in hanging drops

Macromolecule concentration range	0	0	mg/ml
Temperature range	0	0	° C
pH value	4.6	0	
Crystal growth time	0	0	Day

Chemical additions in the reservoir:

Chemical Additive	Range		Unit
ammonium sulfate	2	0	M
Sodium acetate	0.1	0	M

Note: In the above tables, 0 indicates that no data is available.

Reference for Crystal - C2CK.

Reference (R2NE).
Reference Title: "The prosegment-subtilisin BPN' complex:crystal structure of a specific 'foldase'."
Author Names: Gallagher, T; Gilliland, GL; Wang, L; Bryan, P;
Reference Source: Structure 3, 907- 914, 1995.
Reference also for: Crystal C2CJ |

Other Information Related to the Crystal.
Protein Data Bank: 1SPB
NDB:

Fig. 24.4.3.2. A representative example of a crystal entry C2CK for the subtilsin BPN′: prodomain, entry M1MT in the BMCD.

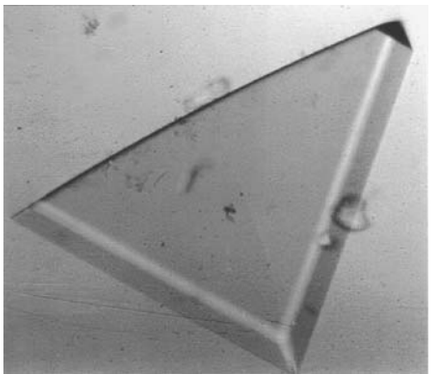

Fig. 24.4.5.1. Crystal of recombinant rat liver glutathione *S*-transferase (Ji *et al.*, 1994) grown from optimized conditions based on the data in BMCD entries M0P3 and C13R.

absence or presence did not affect the crystallization. Crystals of the recombinant enzyme grew within 5 to 10 days (Fig. 24.4.5.1).

24.4.6. Crystallization screens

With the introduction of the fast screen by Jancarik & Kim (1991), almost all attempts to crystallize a protein begin with experiments based on a screen of one form or another. The first screen was based on the ideas put forth by Carter & Carter (1979) in their discussion of the use of incomplete factorial experiments to limit the search, the experience of the investigators themselves, and the experience of others. A number of screens have been developed and even commercialized (*e.g.* Cudney *et al.*, 1994). The first screens were quite general and applicable to a wide range of biological macromolecules, but fast screens based on specific classes of molecules such as RNA soon developed (Scott *et al.*, 1995).

The BMCD is an ideal tool for facilitating the development of screens for general or specific classes of macromolecules. For example, if it were desired to produce a screen for endonucleases, a quick search of the PDB would provide the information in Table 24.4.6.1. An examination of the crystallization conditions of the endonucleases reveals that crystals are grown using a protein concentration ranging from 2.5 to 12.9 mg ml^{-1}, ammonium sulfate, sodium phosphate, or polyethylene glycol 400 to 8000 as precipitants at 4 to 20 °C between pH 4.5 and 8.3. A variety of buffers and standard biochemical additives are also used. From an examination of these parameters, a small subset of the crystallization experiments comprising an endonuclease screen could be developed.

24.4.7. A general crystallization procedure

The use of the BMCD has been incorporated into more general procedures required for the crystallization of a biological macromolecule that has never been crystallized (Gilliland, 1988; Gilliland & Bickham, 1990; Gilliland *et al.*, 1994, 1996). One such general procedure is shown in Fig. 24.4.7.1. The example below illustrates how the data in the BMCD were used to develop this general procedure for soluble proteins. The BMCD can be used to develop analogous procedures for other classes of biological macromolecules. Briefly, in this procedure the purified biological macromolecule is concentrated (if possible) to 10 to 25 mg ml^{-1} and dialysed into 0.005 to 0.025 *M* buffer at a neutral pH or at a pH required to maintain solubility of the biopolymer. Other stabilizing agents such as EDTA and/or dithiothreitol may be included at low concentrations to stabilize the biological macromolecule during the crystallization trials.

Once the protein has been prepared, commercial or customized fast screens are carried out using vapour-diffusion experiments. If crystals are obtained, X-ray diffraction studies are initiated, but frequently small or poor-quality crystals are observed. Experiments that systematically vary the crystallization parameters (pH, ionic strength, temperature *etc.*) are then carried out. Micro- or macroseeding may also be required to optimize crystal growth (McPherson, 1982, 1999).

If the fast screens produce no crystals, a more systematic approach can be undertaken that is based on the data contained in the BMCD. An analysis of the BMCD data reveals that out of the large number of reagents used as precipitating agents, a small set accounts for the majority of the crystals observed. The pH range for all crystals is quite large, but most proteins crystallize between pH 3.0 and 9.0. Even though temperature can be an important factor, crystallization experiments are usually set up at room (~20 °C) or cold-room (6 °C) temperatures. Protein concentration varies quite markedly, but it appears that investigators typically use >10 mg ml^{-1}.

After examining the data in the BMCD, the precipitating agents, ammonium sulfate, polyethylene glycol 8000, 2-methyl-2,4-pentanediol and sodium–potassium phosphate might be selected for the initial crystallization attempts, and experiments might be restricted to a pH range of 3.0 to

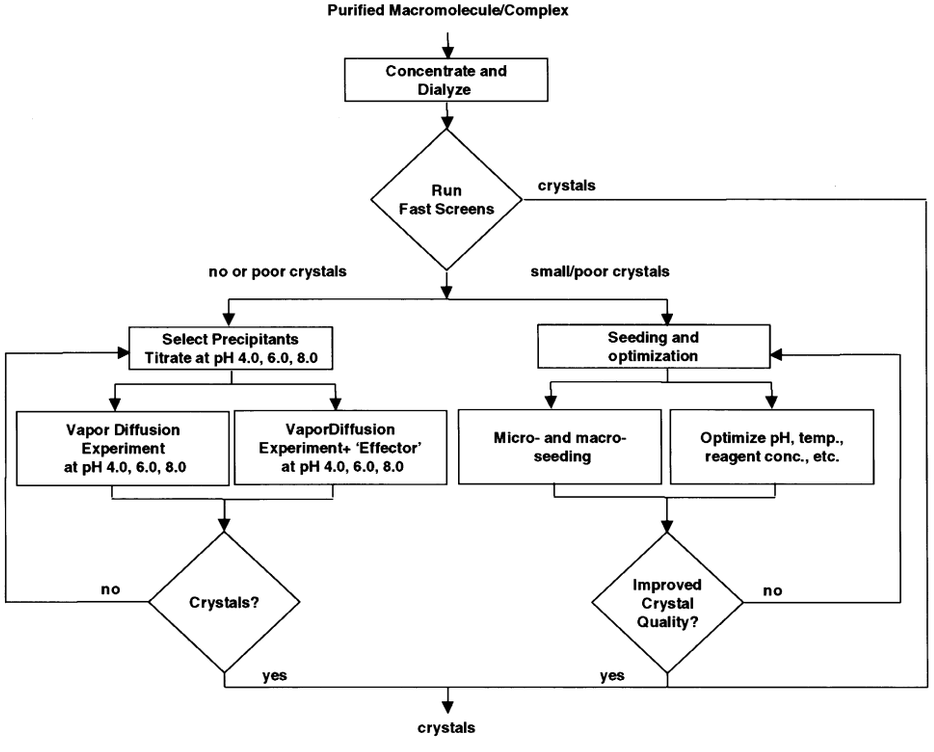

Fig. 24.4.7.1. A general crystallization strategy based on the data contained in the BMCD.

references