INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/642 - 13

College on Methods and Experimental Techniques in Biophysics

28 September - 23 October 1992

Protein Crystallography

C.E. BUGG

University of Alabama, U.S.A.

These are preliminary lecture notes, intended only for distribution to participants.

FAST COMMUNICATIONS

Contributions intended for this section should be submitted to any of the Co-Editors of Acta Crystallographica *or* Journal of Applied Crystallography. *In the letter accompanying the submission authors should state why rapid publication is essential. The paper should not exceed two printed pages (about 2000 words or eight pages of double-spaced typescript including tables and figures) and figures should be clearly lettered. If the paper is available on 5.25" IBM PC compatible or 3.5" Apple Macintosh diskettes it would be helpful if these could be sent with the manuscript together with details of the word-processing package used. Papers not judged suitable for this section will be considered for publication in the appropriate section of* Acta Crystallographica *or in* Journal of Applied Crystallography.

J. Appl. Cryst. (1991). 24, 409–411

Sparse matrix sampling: a screening method for crystallization of proteins

BY JARMILA JANCARIK AND SUNG-HOU KIM

Department of Chemistry and Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720, USA

(*Received* 15 *January* 1991; *accepted* 11 *April* 1991)

Abstract. A set of screening conditions for initial experiments in protein crystallization has been developed, tested, and is herein presented. These solution and precipitant conditions are empirically derived based on known or published crystallization conditions of various proteins in the past, so as to sample as large a range of buffer, pH, additive and precipitant variables as possible, using small amounts of proteins. The 50 crystallization conditions have been tested on 15 previously crystallized proteins, all of which were also crystallized in at least one form by this screen. This method is also shown to be highly successful in the crystallization of proteins which had not previously been crystallized.

Introduction. The most critical step in determining the three-dimensional structure of proteins by the single-crystal X-ray diffraction method is obtaining large single crystals suitable for diffraction studies. Optimal conditions of crystal growth for biological macromolecules in general are very difficult to predict *a priori*. Since the number of variables affecting crystallization, such as concentration, temperature, pH, ionic strength, specific additives and precipitants, is large and combinatorial, the total number of possible solution conditions to be tested is so great as to prohibit an exhaustive search (McPherson, 1982). One approach to overcome this problem is initially to use the incomplete factorial method of Carter & Carter (1979), in which a very coarse matrix of crystallization conditions is explored and the results analyzed to build finer grids around deduced and projected conditions of the initial incomplete factorial. During our experimentation with this method, we realized that we can further minimize the amount of proteins needed for the initial screening by designing a sparse matrix of trial conditions heavily biased on and selected from known or published crystallization conditions. We have chosen three usual categories of parameters as major variables: pH and buffer materials, additives and precipitating agents. For the pH and buffer variables, we have chosen five different pH's: 4.6, 5.6, 6.5, 7.5, 8.5, and for each pH value we have chosen the buffer chemical that has proven to be suitable for protein crystallization in the past. The choice of additives was also based on the past experiences of many laboratories. For the precipitating agent we chose four types: 2-propanol as a volatile agent, 2-methyl-2,4-pentanediol (MPD) and polyethylene glycol (PEG) as nonvolatile agents, various salting out agents, and mixtures of the above. These are listed in Table 1.

Table 1. *Crystallization matrix parameters*

	Precipitating agents		
Non-volatile	Salts	Volatile	Mixture
MPD	Na, K tartrate	2-Propanol	NH_4 sulfate + PEG
PEG 400	NH_4 phosphate		2-Propanol + PEG
PEG 4000	NH_4 sulfate		
PEG 8000	Na acetate		
	Li sulfate		
	Na formate		
	Na, K phosphate		
	Na citrate		
	Mg formate		

pH range: 4.6, 5.6, 6.5, 7.5, 8.5

Salts, additives: Ca chloride, Na citrate, Mg chloride, NH_4 acetate, NH_4 sulfate, Mg acetate, Zn acetate, Ca acetate

By trial and error the resulting multidimensional heavily biased sparse matrix of screening conditions was further simplified by deleting conditions corresponding to many of the matrix elements that can be partially represented by the results of other conditions. The final sparse matrix contains 50 conditions listed in Table 2.

Our experience is that once approximate crystallization conditions are found, it is relatively easy to optimize the

0021-8898/91/040409-03$03.00

Table 2. *Reservoir solutions*

No.	Salt	Buffer	Precipitant
(1)	0.02 *M* Ca chloride	0.1 *M* Acetate	30% MPD
(2)			0.4 *M* Na, K tartrate
(3)			0.4 *M* NH_4 phosphate
(4)		0.1 *M* Tris	2.0 *M* NH_4 sulfate
(5)	0.2 *M* Na citrate	0.1 *M* Hepes	40% MPD
(6)	0.2 *M* Mg chloride	0.1 *M* Tris	30% PEG 4 K
(7)		0.1 *M* Cacodylate	1.4 *M* Na acetate
(8)	0.2 *M* Na citrate	0.1 *M* Cacodylate	30% 2-Propanol
(9)	0.2 *M* NH_4 acetate	0.1 *M* Citrate	30% PEG 4 K
(10)	0.2 *M* NH_4 acetate	0.1 *M* Acetate	30% PEG 4 K
(11)		0.1 *M* Citrate	1.0 *M* NH_4 phosphate
(12)	0.2 *M* Mg chloride	0.1 *M* Hepes	30% 2-Propanol
(13)	0.2 *M* Na citrate	0.1 *M* Tris	30% PEG 400
(14)	0.2 *M* Ca chloride	0.1 *M* Hepes	28% PEG 400
(15)	0.2 *M* NH_4 sulfate	0.1 *M* Cacodylate	30% PEG 8 K
(16)		0.1 *M* Hepes	1.5 *M* Li sulfate
(17)	0.2 *M* Li sulfate	0.1 *M* Tris	30% PEG 4 K
(18)	0.2 *M* Mg acetate	0.1 *M* Cacodylate	20% PEG 8 K
(19)	0.2 *M* NH_4 acetate	0.1 *M* Tris	30% 2-Propanol
(20)	0.2 *M* NH_4 sulfate	0.1 *M* Acetate	25% PEG 4 K
(21)	0.2 *M* Mg acetate	0.1 *M* Cacodylate	30% MPD
(22)	0.2 *M* Na acetate	0.1 *M* Tris	30% PEG 4 K
(23)	0.2 *M* Mg chloride	0.1 *M* Hepes	30% PEG 400
(24)	0.2 *M* Ca chloride	0.1 *M* Acetate	20% 2-Propanol
(25)		0.1 *M* Imidazole	1.0 *M* Na acetate
(26)	0.2 *M* NH_4 acetate	0.1 *M* Citrate	30% MPD
(27)	0.2 *M* Na citrate	0.1 *M* Hepes	20% 2-Propanol
(28)	0.2 *M* Na acetate	0.1 *M* Cacodylate	30% PEG 8 K
(29)		0.1 *M* Hepes	0.8 *M* Na, K tartrate
(30)	0.2 *M* NH_4 sulfate		30% PEG 8 K
(31)	0.2 *M* NH_4 sulfate		30% PEG 4 K
(32)			2.0 *M* NH_4 sulfate
(33)			4.0 *M* Na formate
(34)		0.1 *M* Acetate	2.0 *M* Na formate
(35)		0.1 *M* Hepes	1.6 *M* Na, K phosphate
(36)		0.1 *M* Tris	8% PEG 8 K
(37)		0.1 *M* Acetate	8% PEG 4 K
(38)		0.1 *M* Hepes	1.4 *M* Na citrate
(39)		0.1 *M* Hepes	2% PEG 400, 2.0 *M* NH_4 sulfate
(40)		0.1 *M* Citrate	20% 2-Propanol + 20% PEG 4 K
(41)		0.1 *M* Hepes	10% 2-Propanol + 20% PEG 4 K
(42)	0.05 *M* K phosphate		20% PEG 8 K
(43)			30% PEG 1500
(44)			0.2 *M* Mg formate
(45)	0.2 *M* Zn acetate	0.1 *M* Cacodylate	18% PEG 8 K
(46)	0.2 *M* Ca acetate	0.1 *M* Cacodylate	18% PEG 8 K
(47)		0.1 *M* Acetate	2.0 *M* NH_4 sulfate
(48)		0.1 *M* Tris	2.0 *M* NH_4 sulfate
(49)	1.0 *M* Li sulfate		2% PEG 8 K
(50)	1.0 *M* Li sulfate		15% PEG 8 K

Abbreviations: Tris, 2-amino-2-(hydroxymethyl)-1,3-propanediol; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid. Buffers: Na acetate buffer pH=4.6, Na citrate buffer pH=5.6, Na cacodylate buffer pH=6.5, Na Hepes buffer pH=7.5, Tris.HCl buffer pH=8.5. % is defined as percent by mass.

condition(s) and to finally arrive at the point of obtaining single crystals suitable for the crystallographic studies. Since we, as well as several other laboratories, have had a substantially higher rate of success with this approach than with more systematic approaches used in the past, we would like to share our experience with others.

Table 3. *Proteins that have been previously crystallized*

All the proteins were obtained commercially (except those marked with *) and used without further purification. Natural cofactors were present whenever needed.

Protein (source), M_r (Da)	Crystallization conditions	Crystal size (largest dimension) (mm)
Lysozyme (chicken egg white, Sigma), 14 400	9, 13, 15, 28, 32	1.0
Myoglobin (sperm whale, Sigma), 17 200	1, 5, 13	0.8
Trypsin (bovine pancreas, type 3, Sigma), 24 000	3, 17, 30	0.8
Hemoglobin (horse blood, Sigma), 64 500	14, 27, 30	1.0
Pepsin (porcine stomach mucosa, Sigma), 35 000	2, 6, 28	0.5
Ribonuclease A (bovine pancreas), 13 700	6, 11, 13, 23	0.7
Catalase (bovine liver, Sigma), 247 500	14, 15, 22, 27, 30	0.8
Papain (papaya latex, Sigma), 23 000	2, 6, 24	1.0
Insulin (porcine pancreas, Sigma), 6000	13, 26	0.6
Thrombin (bovine plasma, Sigma), 33 500	11, 32	0.5
Ribonuclease B (bovine pancreas, Sigma), 13 000	17, 22, 37	0.7
Tropomyosin (chicken muscle, Sigma), 36 000	4 , 22, 38	0.5
Tropomyosin (rabbit muscle), 37 000	4, 32, 38	0.7
Ubiquitin (bovine red blood cells, Sigma), 8500	6, 2, 37	0.4
*Eco*RI endonuclease,* 31 000	14, 15, 24	0.3
*Eco*RI-DNA (GAATTC) complex,* 33 000	4, 15	1.0
Monellin (Sigma), 10 500	4, 14, 19, 31	1.5

We have tested 15 previously crystallized proteins (Table 3) using this technique, and we have found each of them crystallized from several different conditions in our matrix, one of which being the same or similar to the reported crystallization condition. We have also used this method on many proteins which have not previously been crystallized, and our rate of success is quite high. Proteins which have failed to crystallize using this technique have been shown to be non-homogeneous by analytical techniques such as gel electrophoresis and isoelectric focusing. This type of screening approach is simple and practical for finding initial crystallization conditions of proteins when the amount is so limited that more comprehensive screening methods are impossible or impractical.

Material and methods. After preparing 50 different solutions, each containing the appropriate additive, buffer and precipitant as indicated in Table 2, we use the hanging-drop method [see, for example, McPherson (1976)] to screen for protein crystallization conditions. This method involves suspending 2–4 μl droplets over a 0.7 ml reservoir. The droplet contains 1–2 μl of aqueous protein solution and 1–2 μl of the reservoir solution. We found that the most favorable concentration of the protein for initial tests is about 10 mg ml^{-1}. The concentration of the precipitant in the reservoir is twice that in the droplet, giving

a reasonably slow rate of vapor-phase equilibration for dehydration of the droplet or precipitation of the protein. All of the solutions are made with doubly distilled water and are passed through 0.45 μm millipore filters. Crystallization experiments were done in Linbro plastic plates (Linbro model 76-033-05) originally designed for cell culture, covered with 22 × 22 mm cover slips which have been siliconized with 1% Prosil-28 (PCR Inc., Gainsville, Florida). Two sets of experiments are prepared simultaneously, one at room temperature and a second at 277 K. The droplets are carefully examined under a stereoscopic microscope every day during the first several weeks and then on a monthly basis. Quite often the initial crystals are very small, and therefore it is essential to examine each drop very carefully, scanning the focal plane of the microscope at all levels of the droplet.

Results and discussion. Results of this crystallization method are shown in Tables 3 and 4. Table 3 illustrates that 100% of the proteins which have been previously crystallized were also crystallized under many different conditions using our method. We have also had high rates of success in crystallization of new proteins (Table 4). Our results indicate that when we are able to crystallize proteins using this method, they usually crystallize in several different conditions and, in many cases, are polymorphic. Of the conditions in which initial crystallization is observed, one form is usually favored by size or morphology for optimization.

The main advantage of this method is that we can quickly test wide ranges in pH, salts and precipitants using a relatively small amount of sample. The initial screening of 50 widely varying conditions requires approximately 1–2 mg of protein. Furthermore, because the same conditions are used for initial screening of all protein samples, this approach is well suited for automation. We intend to continuously update and improve the conditions of the test matrix as the results of crystallization of new proteins become available. Any data or information in this respect would be greatly appreciated.

The authors are especially appreciative of the contributions of Dr Marcos Hatada in the initial planning and testing of crystallization conditions and for his continuing advice. This work was supported by grants from National Institute of Health (CA45593, DC45593) and Department of Energy (contract No. DE-AC03-76SF00098).

Table 4. Tests on proteins not previously crystallized

None of the proteins listed are commercially obtained.

Protein, M_r (Da), cofactor	Crystallization conditions	Crystal size (largest dimension) (mm)
Thaumatin, 21 000	2, 14, 15, 28, 33	1.5
Allophycocyanin, 16 000	7, 12, 13, 16, 19, 28	0.2
Phytochrome, 124 000	—	—
*Eco*RI methylase, 39 000	—	—
Epidermal growth factor, 5800	—	—
Ferric uptake regulator protein fragment, 17 000	—	—
c-H-ras protein (Gly 12), 21 000, GDP	12	0.7
c-H-ras protein (Val 12), 21 000, GDP	12	0.8
c-H-ras protein (Leu 61), 21 000, GDP	12	0.7
c-H-ras protein (Gly 12, 1–171), 19 000, GDP	6, 13	1.5
c-H-ras protein (Val 12, 1–171), 19 000, GDP	6, 13	1.5
c-H-ras protein (Leu 61, 1–171), 19 000, GDP	6, 13	1.5
c-H-ras protein (Gly 12, 1–171), 19 000, GDPCP	18, 28	1.5
c-H-ras protein (Val 12, 1–171), 19 000, GDPCP	18, 28	1.0
Single-chain monellin 1, 9 500	15, 36	0.7
Single-chain monellin 4, 9500	4, 15, 19, 22, 29	0.3
Single-chain monellin 5, 9500	9, 20, 31, 32, 44	0.4
Single-chain monellin 6, 9500	—	—
Single-chain monellin 7, 9500	14, 19, 30	0.2
Single-chain monellin 12, 9500	—	—
Formyltetrahydrofolate synthetase, 230 000	7, 11, 25	0.1
Tumor necrosis factor, 17 000	10, 15, 20, 26, 36	1.2
Human M-CSF, truncated, 34 000	9, 10, 14, 15, 20, 23	1.5
Bacteriorhodopsin, 27 000	32	0.1
Methallothionein (human), 6500, Zn	6, 12, 19	0.1
Methallothionein (human), 6500, Cd	6, 24	0.1
Methallothionein (horse kidney), 6500, Cd	33, 34	0.7
'Tar' receptor, periplasmic domain, 22 000	25	0.2
His J protein, 26 127	6	1.0
Lao (Lys-Arg-Orn) binding protein, 26 000	28	1.2
Lao protein–Arg complex, 26 000	22	1.9

References

CARTER, C. W. JR & CARTER, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.

MCPHERSON, A. JR (1976). *Methods Biochem. Anal.* **23**, 249–345.

MCPHERSON, A. JR (1982). *Preparation and Analysis of Protein Crystals*. Wiley: New York.