ISSN: 0976-3104
REGULAR ISSUE
Gaur

THE IIOAB JOURNAL

ARTICLE OPEN ACCESS

ESTIMATION OF PEG TYPES AND THEIR CONCENTRATION DURING PROTEIN CRYSTALLIZATION

Rajneesh K. Gaur
Department of Biotechnology, Ministry of Science and Technology, 814, Block-2, C.G.O. Complex, Lodhi Road, New Delhi -110003, INDIA

ABSTRACT

Polyethylene Glycol (PEG) is a major precipitant in protein crystallization. This study focused on analytical estimation of different PEG types and their concentration on protein crystallization. The results indicate that ~84% of soluble proteins and ~78% of membrane proteins are crystallized with six PEG types (PEG3.35k>4k>8k>0.4k>6k>2kMME) and four PEG types (PEG3.35k>4k>6k>8k) respectively. The ~48% of soluble and ~62% of membrane proteins are crystallized with PEG3.35k & PEG4k only. Therefore, PEG4k may be used as an independent screening agent in PEG only commercial screens. Except PEG0.4k, remaining five PEG types contribute ~15-20% of soluble protein crystallization at 25% w/v concentration. The various classes of soluble protein i.e. All Alpha, All Beta and Alpha & Beta (a/b & a+b) does not show distinct preference for different PEG types. These results can be used to improve the PEG based protein crystallization commercial screens.

Received on: 27th-May-2016
Revised on: 08th-June-2016
Accepted on: 08th– June-2016
Published on: 08th– July-2016

KEY WORDS

Polyethylene Glycol (PEG) type, Polyethylene Glycol (PEG) concentration, Protein Crystallization

***Corresponding author:** **Email:** meetgaur@gmail.com; **Tel:** 011-24360295 **Fax:** 011-24360295

INTRODUCTION

Protein crystallization is a complex process, which is influenced by multiple parameters [1]. Empirical knowledge facilitated development of several screening methods to enhance the overall crystallization hits [2]. Efforts are continuing to improve the protein crystallization screens [3, 4]. Despite the remarkable progress made, protein crystallization is still a major bottleneck [5].

The most successful precipitants for protein crystallization are Polyethylene Glycol (PEG) and Ammonium Sulfate [6]. PEG is widely used in the crystallization of proteins, which are either crystallized alone or in complex. The various PEG related parameters such as different PEGs types, their concentration and molecular mass as well as influence of pH and salt concentration of the crystallizing solution have been studied for improving the crystallization process and introducing new commercial screens [7 ,8]. However, the influence of different PEG type and their concentration on crystallization of major group of proteins and their classes needs further exploration. With increasing number of structures available in the Protein Data Bank (PDB; till date ~98720 proteins structures solved through X-ray), the influence of different PEG types & concentration on protein crystallization can be easily ascertained.

METHODOS

Protein sequences having 30% sequence identity and crystallized with PEG are downloaded from Protein Data Bank (PDB) [9]. Out of the downloaded 9410 X-ray diffracted protein entries; two separate experimental dataset comprised of 1441 soluble protein and 43 membrane protein entries has been manually curated. The protein entries are curated after excluding the entries crystallized in complex with any type of ligand including protein/peptide/any chemical entity such as ATP etc. and also the entries possess inadequate and insufficient crystallographic information. The experimental dataset for soluble proteins is further divided for analytical purpose into four sub-datasets of 'All Alpha (423)', 'All Beta (323)', 'Alpha and Beta [a/b (165); a+b (530)]' proteins as per the Structural

THE IIOAB JOURNAL
ISSN: 0976-3104

Classification of Protein (SCOP) [10]. The percentage of proteins crystallized by different PEG types & their concentration profile was manually calculated and analyzed.

RESULTS AND DISCUSSION

PEG is known as one of the main precipitant for protein crystallization. PEG parameters directly influence the protein crystallization is determined either through data mining [11] or surveys [12]. Considering the tremendous growth of PDB database, two experimental datasets of protein (soluble and membrane) entries having 30% sequence identity and crystallized with PEG is prepared to determine the relationship between PEG types and their concentration on the crystallization of Proteins. The dataset of soluble proteins is further subdivided in to various classes of proteins to assess the occurrence of preference(s) towards PEG types and their concentration by any particular class. Considering the statistically less number of entries, the membrane protein dataset is not subdivided in to various classes. The 30% sequence identity criteria selected to incorporate only the unique sequences or sequences of least similarity in the experimental dataset [13]. The protein-protein complexes crystallized with PEG is not included in the dataset as they are studied earlier in detail [12]. Both the datasets were checked manually and found to possess the non-redundant crystallization conditions. In addition to PEG as a precipitant, ~20% of experimental dataset entries also possess Ammonium Sulfate or Sodium Chloride or MPD or combination of them as an additive.

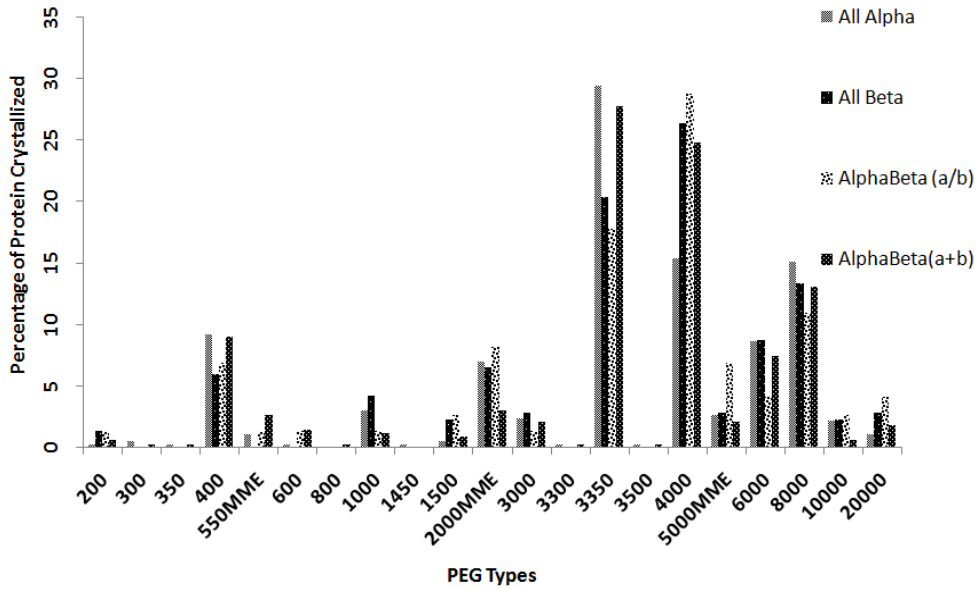

Fig: 1 shows the percentage of various classes of soluble proteins (alone) crystallized with different PEGs. The classes of protein are All Alpha (Grey), All Beta (Black) and Alpha & Beta (a/b - Starred & a+b – Black and White points)) types. The PEG types used based on the crystallizing conditions reported in PDB database.

In total, ~84% of soluble proteins are crystallized by just six PEG types [PEG3.35k (26.08%) > PEG4k (21.95%) > PEG8k (13.8%) > PEG0.4K (8.26%) > PEG6k (7.96%) > PEG2kMME (5.64%)] [**Figure-1**]. These results are in agreement with earlier studies, which shows that most of the proteins are crystallized with few PEG types mainly 3.35k, 4k, 6k & 8k [14, 15]. The results obtained here imply that only six PEG types substantially cover protein crystallization space. The percentage of soluble proteins (~62%) crystallized with three PEG types i.e. PEG3.35k, 4k & 8k shows that nearly double the percentage of proteins crystallized with these three PEG types in comparison to earlier studies. The difference is due to the inclusion of proteins having unique sequences or sequences of least similarity alone in this study. Out of the six PEG types, the two PEGs i.e. PEG3.35k and PEG4k contribute in the crystallization of ~48% soluble proteins. PEG3.35K used alone as a screening agent in commercial PEG screens such as PEG/ion screen (Hampton Research) and PEG suite (Qiagen). These results suggest that PEG4k, like PEG3.35k, can also be used independently as a screening agent for preliminary screening of proteins in commercial crystallization screens especially when the protein sample is precious. The commercial availability of an independent PEG3.35k & PEG4k screen will certainly enhance the overall efficiency of protein crystallization.

THE IIOAB JOURNAL
ISSN: 0976-3104

In total, ~78% of the membrane proteins are crystallized by just four PEG types [PEG3.35k (35.1%) > PEG4k (27.0%) > PEG6k (8.1%) > PEG8k (8.1%)] **[Figure-2]**. Like soluble proteins, PEG3.35k and PEG4k also contributes remarkably in the crystallization of membrane proteins (~62%). These results are in contrast to a study where PEG4k is reported as most successful PEG precipitant [16]. The difference is due to the inclusion of all the membrane proteins in this study rather than only the outer membrane proteins. The crystallization of various membrane protein classes is not studied as a result of less number of entries.

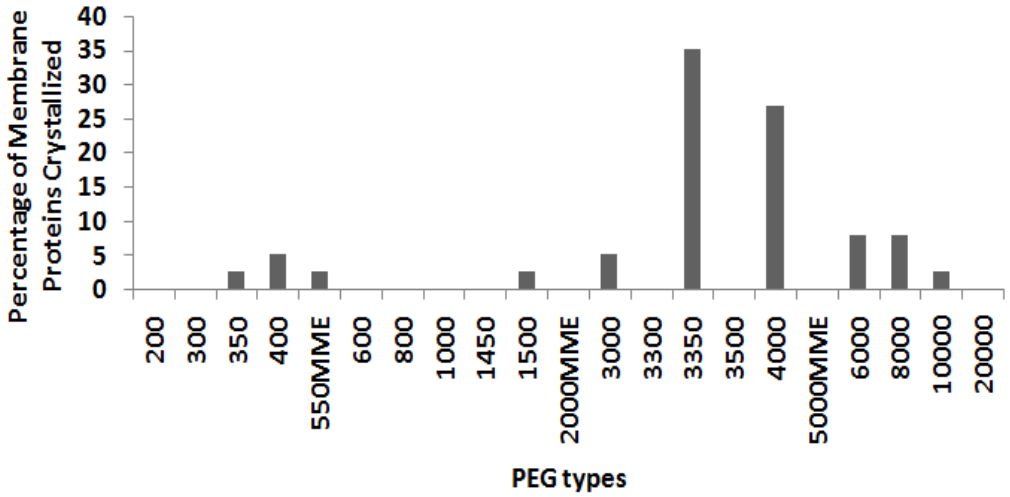

Fig: 2 Shows the percentage of membrane proteins (alone) crystallized with different PEGs. The PEG types used based on the crystallizing conditions reported in PDB database.

To determine the occurrences of preference(s) of various classes of soluble proteins towards six PEG types, the experimental dataset of soluble proteins is divided into sub-dataset of 'All Alpha, All Beta, Alpha and Beta (a/b and a+b) proteins as per the SCOP classification. It is revealed that the four classes of soluble proteins are crystallized with six PEG types with minor variation **[Figure-1]**. The PEG types i.e. PEG3.35k and PEG4k cumulatively leads to the crystallization of All Alpha (44.87%), All Beta (46.76%), Alpha and Beta [(a/b - 46.58%) & (a+b – 52.54%)] proteins, which is near to the total value (~48%) observed for the soluble proteins. 'Alpha and Beta (a/b)' proteins (17.81%) show ~9% (from the total) less percentage crystallization with PEG3.35k in comparison to other protein classes. Similarly, 'All Alpha' proteins (15.41%) show ~6% (from the total) less percentage crystallization with PEG4k in comparison to other protein classes. Furthermore, the percentage of 'Alpha and Beta (a/b)' protein crystallized with PEG6k is nearly half as compared to other protein types, while crystallized two and half times more with PEG5kMME. Similarly, the percentage of 'Alpha and Beta (a+b)' proteins crystallized with PEG2kMME is nearly one third as compared to 'Alpha and Beta (a/b)' protein type. These differences can be attributed to the influence of secondary structure composition on protein packing [17].

The data is further analyzed to determine the concentration profile of six PEG types resulted in the highest percentage of soluble protein crystallization **[Figure-3]**. The concentration profile of six PEG types shows distinct concentration peaks. PEG0.4K shows peaks at 2% & 35% v/v (31.88% of total protein crystallized); PEG2k at 15%, 20%, 25% & 30% w/v (49.12% of total protein crystallized); PEG3.35k at 20% & 25% w/v (60.31% of total protein crystallized); PEG4k at 25%, 30% & 35% w/v (43.81% of total protein crystallized); PEG6k at 15% & 25% (46.15% of total protein crystallized) and PEG8k at 25% w/v (21.04% of total protein crystallized) respectively. Similar studies regarding the PEG concentration and their relative frequency are also carried out by Peat etal. 2005 [11]. The 'relative frequency' was the count of the number of occurrences of a PEG at a given concentration divided by the total number of conditions containing that PEG. Their results indicate that the Low molecular weight (LMW) PEGs show no clear preferred concentration. However, our results clearly show a sharp peak for percentage of proteins crystallized at 2% v/v for PEG0.4k, which suggests that PEG0.4k acts as a precipitant as well as an additive at low concentration. Further, these results justify the usage of PEG0.4k in the concentration range of 30-35% in commercial screens. Peat et al., 2005 [11] reported that Medium Molecular Weight (MMW) PEGs (1k-8k) shows their concentration peak between 20-25%, however, our results shows wider range for MMW PEGs. In our study, it is observed that MMW PEGs facilitate crystallization of ~15-20% of protein at 25% w/v concentration.

THE IIOAB JOURNAL
ISSN: 0976-3104

Furthermore, the four PEG types also shows other major concentration (% w/v) peaks i.e. PEG2kMME (15%, 20% & 30%), PEG3.35k (20%), PEG4k (30% & 35%), PEG6k (15%). The doubling of protein crystallization percentage with PEG2k from 20% to 25% w/v is due to higher volume exclusion effect leading to more protein aggregation. The concentration profile for membrane proteins is not mapped due to less number of entries in the dataset. PEG3.35k & 4k facilitate at 20% & 30% w/v concentration the crystallization of Membrane proteins.

The occurrence of few concentration peaks for six PEG types complies with the concentration range found in most of the commercial kits. However, currently none of the commercially available PEG based screens possess this combination of six PEG types and the concentrations revealed out of this study. PEG suite (Qiagen) is the only screen using various PEG types at 25% w/v concentration, as observed in our results. These results indicates that five PEGs (i.e.2k, 3.35k, 4k, 6k & 8k) can be incorporated at the concentrations, observed out of this analysis, in commercial screens for augmenting the chances of crystallization.

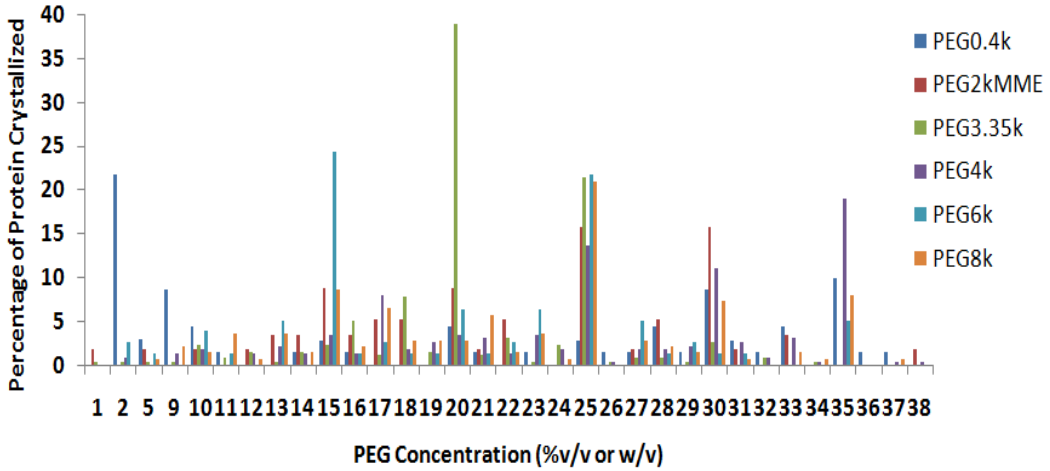

Fig:3 shows the percentage of different PEG concentrations and percentage of proteins crystallized. The six PEG types, whose concentration plotted, are PEG0.4k (dark blue), PEG2kMME (Red), PEG3.35k (Green), PEG4k (violet), PEG6k (light blue) and PEG8k (organge).

In this study, no clear correlation observed between the concentration profile and molecular weight of different PEGs as opposed to a generalized relationship of decreasing the median PEG concentration with increasing molecular weight of PEGs as reported for protein-protein complexes [12]. This type of diffused pattern may be due to the effect of multiple parameters such as viscosity, interplay of attractive and repulsive forces, ionic strength etc. on protein crystallization [18].

SYNOPSIS

Medium molecular weight (2k-8k) PEG types contribute significant percentage of protein crystallization. A major fraction (~50-60%) of proteins is crystallized with only PEG3.35k & PEG4k types. The results obtained can be used to improve the commercial crystallization screens.

CONCLUSION

In conclusion, a similar study as carried out earlier by Fazio et al., 2014 [15] is desired to evaluate the success of PEG based screens in crystallization space to validate the results obtained. Alternatively, a PEG screen prepared on the basis of the results obtained in this study may be used in high throughput protein crystallization for validating the success of a new recipe for improving the efficiency of protein crystallization.

THE IIOAB JOURNAL
ISSN: 0976-3104

CONFLICT OF INTEREST

There is no conflict of interest.

ACKNOWLEDGEMENT

None.

FINANCIAL DISCLOSURE

No financial support was received for this study.

REFERENCES

[1] Kluser Rosenberger F. [1996] Protein Crystallization *J Cryst Growth* 166: 40-54.
[2] Shieh HS, Stallings WC, Stevens AM, Stegeman RA. [1995] Using sampling techniques in protein crystallization Acta Cryst D51: 305-310.
[3] Rupp B, Wang J. [2004] Predictive models for protein crystallization Methods 34: 390-407.
[4] Gorrec F. [2009] The MORPHEUS protein crystallization screen *J Appl Cryst* 42: 1035-1042.
[5] Luft JR, Newman J, Snell EH. [2014]. Crystallization screening: the influence of history on current practice Acta Cryst F, *Struc Biol Commun* 70: 835-853.
[6] Dumetz AC, Chockla AM, Kaler EW, Lenhoff AM. [2009] Comparative Effects of Salt, Organic, and Polymer Precipitants on Protein Phase Behavior and Implications for Vapor Diffusion Growth Des 9: 682-691.
[7] Page R, Stevens RC. [2004] Crystallization data mining in structural genomics: using positive and negative results to optimize protein crystallization screens Methods 34: 373-389.
[8] Newman J, Egan D, Walter TS, Meged R, Berry I, Jelloul MB, Sussman JE, Stuart DI, Perrakis A. [2005] Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy Acta Cryst D61: 1426-1431.
[9] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. [2000] The protein Data *Bank Nucleic Acids Res* 28: 235–242.
[10] Murzin AG, Brenner SE, Hubbard T, Chothia C. [1995] SCOP: a structural classification of proteins database for the investigation of sequences and structures *J Mol Biol* 247: 536-540.
[11] Peat TS, Christopher JA, Newman J. [2005] Tapping the Protein Data Bank for crystallization Information Acta Cryst D61: 1662-1669.
[12] Radaev S, Sean Li, Sun PD. [2006] A survey of protein-protein complex crystallizations Acta Cryst D 62: 605-612.
[13] Pearson WR. [2013] An Introduction to Sequence Similarity ("Homology") Searching Current Protocol in Bioinformatics, 42:3.1.1-3.1.8.
[14] Li C, Kirkwood KL, Brayer GD. [2007] The Biological Crystallization Resource: Facilitating Knowledge-Based Protein Crystallizations Crystal Growth Des 7: 2147-2152.
[15] Fazio VJ, Peat TS, Newman J. [2014] A drunken search in crystallization space Acta Cryst F 70: 1303-1311.
[16] Newstead S, Hobbs J, Jordan D, Carpenter EP, Iwata S. [2008] Insights into outer membrane protein crystallisation Mol Membr Biol 25: 631-638.
[17] Fleming PJ, Richards FM. [2000] Protein packing: dependence on protein size, secondary structure and amino acid composition ,*J Mol Biol* 299: 487-498.
[18] Kozer N, Kuttner YY, Haran G, Schreiber G. [2007] Protein-Protein Association in Polymer Solutions: From Dilute to Semidilute to Concentrated *Biophysical J* 92: 2139-2149.

ABOUT AUTHOR

Dr. Rajneesh Kumar Gaur *is currently associated with Department of Biotechnology, Ministry of Science and Technology, New Delhi, India. His interest is the analyses of biological data available in public domain databases.*