ARTICLE

AMMONIUM SULPHATE CONCENTRATION OPTIMIZATION AND ITS RELATION WITH PROTEIN PARAMETRS FOR CRYSTALLIZATION

Rajneesh K. Gaur

Dept. of Biotechnology, Ministry of Science and Technology, CGO complex, Lodhi Road, New Delhi, INDIA

ABSTRACT

Ammonium sulphate (AS) is the second most utilized precipitant in protein crystallization. This study focused on determining the optimum AS concentration range for crystallizing the four classes of proteins and the relation between the theoretical protein parameters such as iso-electric point and aliphatic index and AS concentration. The data analysis indicates that the AS concentration in 1.5M-2.5M range leads to crystallization of 61.83% of single and soluble proteins and nearly 57% of proteins crystallized as complex structures. In this range, the four classes of proteins show 65.19% (All Alpha), 63.02% (All Beta), 61.09% (Alpha and Beta, α+β) and 54.27% (Alpha and Beta, α/β) crystallization respectively. There is an inverse relation between theoretical iso-electric Point (pI) and AS conc. facilitate crystallization of 'All Alpha' and 'All Beta' proteins and direct correlation for 'Alpha and Beta' Proteins. It is further observed that there is an inverse relationship between aliphatic index of a protein and AS conc. facilitates crystallization of 'All Alpha' and 'Alpha and Beta' proteins, while correlation is direct for 'All Beta' proteins. These results can be used to improve the existing commercial crystallization screens as well as to predict AS conc. to facilitate crystallization of proteins based on their theoretical isoelectric point and aliphatic index. In conclusion, the optimum ammonium sulphate concentration for crystallization of four classes of proteins is unknown. The data analysis revealed that, in general, ~62% of proteins are crystallized with 1.5M-2.5M of ammonium sulphate concentration range and these results can be used to improve the commercial crystallization screens.

INTRODUCTION

Protein crystallization is a complex phenomenon. In general, protein crystallization is dependent on variety of factors such as precipitant conc. (used either at saturation or in molar quantity), buffer pH & ionic strength and several protein based parameters such as its solubility, isoelectric point, molecular mass, hydropathy & aliphatic index, etc. [1]. The most successful precipitants for protein crystallization are Polyethylene Glycol (PEG) and Ammonium Sulfate (AS) [2].

The available commercial screens cover sufficient crystallization space while accommodating the precipitant related parameters, buffer PH and salt conc. etc. [3, 4]. In general, number of commercial screens is available; there efficiency of protein crystallization needs improvement through data mining [5]. As a result of statistical analysis, a new crystallization screen called as 'Berkley Screen' is recently available [6].

The data analysis for estimation of PEG types and their conc. in crystallization of various class of protein has been recently reported [7]. Similar study is required for AS as precipitant considering the large number of X-ray based structures is available in the Protein Data Bank (PDB; till date ~124337 protein structures in total). This study is based on data analysis to determine the influence of AS concentration on different classes of single & soluble protein crystallization and the related protein parameters. The outcome of the study will be helpful in improving the efficiency of available AS crystallization screens or formulating new screens as well as to predict the AS conc. for crystallizing a particular class of protein utilizing the theoretical parameters of a protein sequence.

METHODS

The soluble proteins crystallized with Ammonium Sulphate (AS) and having 30% sequence identity are downloaded from Protein Data Bank (PDB) [8]. Out of the 1062 downloaded X-ray diffracted protein entries, only 162 protein entries are used in experimental dataset. The number of membrane protein entries is insufficient for data analysis. 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, FAD etc. and those possess inadequate and insufficient crystallographic information. Only the non-redundant crystallization conditions were incorporated in the experimental dataset. For analytical purpose, the experimental dataset of soluble proteins is further divided into four sub-datasets of 'All Alpha (28)', 'All Beta (50)', 'Alpha and Beta [a/b (39); a+b (45)]' proteins as per the Structural Classification of Protein (SCOP) [9]. The percentage of proteins crystallized at a particular Ammonium Sulphate Concentration is manually calculated. The theoretical protein parameters i.e. isoelectric point and Aliphatic index are calculated by using 'Prot Param' tool available on Expasy server [10] and the entire data is analyzed manually.

Corresponding Author
Email:
meetgaur@gmail.com
Tel.: +11-24362982
Fax: +11-24360745

RESULTS AND DISCUSSION

Ammonium Sulphate (AS) and PEG are the two main precipitant used for protein crystallization. This study is focused on AS concentration (M) determination facilitating the maximum percentage of various classes of soluble protein crystallization and its relation with two theoretical protein parameters i.e. iso electric point & Aliphatic index. Earlier also the AS concentration optimization for protein crystallization has been reported [11].
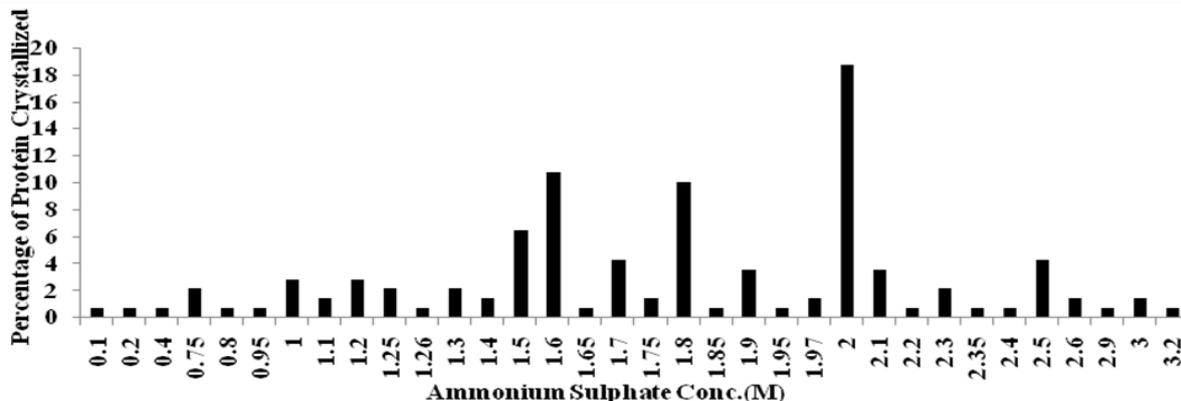

Fig. 1: Shows the percentage of soluble proteins crystallized with different Ammonium Sulphate Concentrations (M).
....

In last few years, there is a tremendous growth in PDB database of protein structures crystallized through X-ray diffraction method, therefore, it is pertinent to determine the optimized AS concentration for protein of various classes. As a result, an experimental dataset of soluble single proteins having 30% sequence identity was prepared and subdivided in to four subsets i.e. 'All Alpha', 'All Beta', Alpha and Beta (alpha/beta & alpha+ Beta). The manual analysis of the overall dataset revealed that the AS concentration in decreasing order of 2M>1.6M>1.8M>1.5M facilitated the percentage crystallization of soluble proteins [Fig. 1]. These four AS concentrations leads to the crystallization of 46.03% of proteins in total and 18.7%>10.79%>10.07%>6.47% independently. In addition, the AS concentration each of 1.7M, 1.9M, 2.1M & 2.5M results in crystallization of 4.31%, 3.59%, 3.59% & 4.31% of proteins independently, which cumulatively leads to 15.8% of protein in total. Therefore, the range of AS concentration leading to crystallization of 61.83% of single soluble proteins is 1.5M-2.5M. In market, number of commercial kits is available such as Ammonium Sulphate suite (Qiagen, Germany) & Grid Screen AS (Hampton Research, USA). In these commercially available screens, the AS concentration used either in multiple of a particular concentration such as multiples of 0.8M AS is used in Grid Screens (Hampton Research) or use of a particular concentration in majority of conditions such as 2.2M in Ammonium Sulphate Suite (Qiagen). In contrast to the deduced AS Conc. range, the existing commercial screens using extremes of AS Conc. However, the AS conc. range of 1.5M to 2.5M if included at an interval of 0.2M in available commercial AS screens might enhance their efficiency.

Table 1. Shows the Ammonium Sulphate (AS) Concentration (M) and percentage of four classes of single and soluble protein crystallized at 5% or above AS conc.

| Ammonium Sulphate concentration (M) | Protein Classes (as per SCOP classification) | | | |
| | All Alpha | All Beta | Alpha and Beta | |
			(α/β)	(α+β)
2.5	-	5.16	5.71	-
2.0	21.73	15.78	20.0	22.22
2.1	8.69	7.89	-	-
1.9	-	-	5.71	5.55
1.8	17.39	7.89	8.57	11.11
1.7	-	5.26	-	5.55
1.6	8.69	13.15	14.28	8.33
1.5	8.69	7.89	-	8.33
1.3	-	5.26	-	-
1.2	-	5.26	-	-
1.1	-	-	-	5.55
1.0	-	-	5.71	-
0.75	8.69	-	-	-

BIOINFORMATICS

Further, the AS concentration resulting in the crystallization of four classes of proteins crystallized was studied [Fig. 2].The results show that all the four protein classes shows the maximum crystallization percentage at three AS conc of 2.0M>1.8M>1.6M [Table 1]. All the protein classes show highest percentage of crystallization at 2.0M AS conc. However, 'All Alpha' and 'Alpha and Beta (α+β)' shows higher percentage of crystallization at 1.8M AS conc. in comparison to 'All Beta' and 'Alpha & Beta (α/β)' proteins. These classes show higher crystallization percentage at 1.6M AS conc. in comparison to 'All Alpha' and 'Alpha and Beta (α+β)' class of proteins. These four classes of proteins show 36-47% of crystallization percentage at three AS conc. i.e. 2.0M, 1.8M and 1.6M. 'Alpha and Beta (α/β & α+β)' protein class show similar percentage of crystallization i.e. ~42%, while 'All Alpha (47.81%)' & 'All Beta (36.82%)' proteins remains at two extreme at these three AS conc. It may be due to the fact that the alpha-helix structure is determined by tertiary structure and beta sheets by intrinsic properties of the residues in the strand [12]. In alpha helices, the tertiary structure formation involves variation of amino acid residues, while beta sheet residues are more conserved [13]. Beta sheets are stabilized by hydrophobic contacts and backbone hydrogen bonding. Alpha helices are largely stabilized by backbone H-bonding i.e. local interactions dominate in a Helix, whereas a sheet is stabilized by long range contacts. So, any disturbance of hydrophobic bonds in Beta sheets could result in exposure of hydrophobic residues and leading to difficulty in crystallization. Therefore, low value for 'All Beta' proteins. Furthermore, though the residues in beta sheets are conserved, possibly the existence of high mobility of folds or other inconsistent secondary structures is not allowing the crystallization of high percentage of 'All Beta' proteins, while the alpha-helices in 'All-Alpha' proteins possess greater flexibility to accommodate such disturbances due to better protein-protein/water interaction and leading to high percentage of protein crystallization of 'All-Alpha' proteins at these three AS conc. The observed difference in AS conc. in 'Alpha and Beta' class of protein crystallization is due to the proportional difference of Helices & Beta Sheets in this class of protein, which leading to various levels of repulsive protein-protein interactions [14].
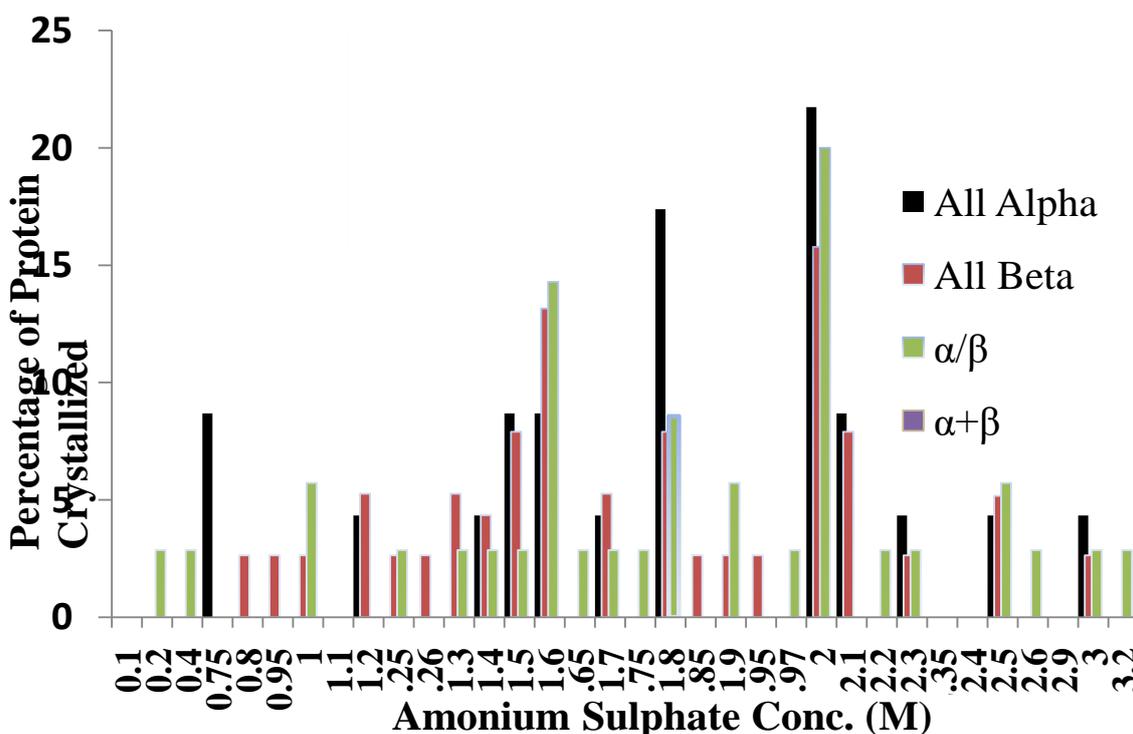

Fig. 2: Shows the percentage of various classes of soluble proteins crystallized with different Ammonium Sulphate Concentrations (M). The different classes of proteins are All Alpha (Black), All Beta (Maroon), Alpha and Beta (α/β – Green & α+β – Purple).

....

An interesting observation is that four classes of proteins show 65.19% (All Alpha), 63.02% (All Beta), 61.09% (Alpha and Beta, α+β) and 54.27% (Alpha and Beta, α/β) percentage crystallization respectively within the AS conc. range of 1.5M-2.5M [Fig. 2]. Exceptionally, 'All Alpha (8.69%)' proteins also show good chances of crystallization at 0.75M. These results suggest significant percentage of the four classes of proteins is crystallized within a narrow AS conc. range of 1.5M-2.5M. Therefore, the narrow AS conc. range can be utilized to formulate new commercial screens for minimizing the wastage of precious protein samples and to obtain the crystallization conditions quickly. In addition, the new screens can also accommodate a separate slot at 0.75M AS conc. to enhance the crystallization efficiency for 'All Alpha' class of proteins. High throughput platforms or Structural Consortiums provide ample scope to experimentally validate these results for single and soluble proteins, whose structures are not yet available.

Table 2. The Ammonium Sulphate (AS) Concentration (M) and percentage of four classes of protein crystallized at 5% or above AS conc. as complex

| Ammonium Sulphate concentration (M) | Protein Classes (as per SCOP classification) | | | |
| | All Alpha | All Beta | Alpha and Beta | |
			(α/β)	(α+β)
3.5	6.66	9.52	-	-
3.2	6.66	-	-	-
2.9	6.66	-	-	-
2.8	-	-	-	9.52
2.5	13.33	-	-	-
2.4	-	-	5.26	-
2.2	13.33	-	-	-
2.1	-	-	5.26	9.52
2.0	6.66	33.33	21.05	23.80
1.9	-	-	7.89	-
1.8	6.66	19.04	5.26	14.28
1.7	13.33	-	-	-
1.6	6.66	-	7.89	-
1.5	-	-	10.52	19.04
1.4	6.66	-	5.26	-
1.3	-	-	5.26	-
1.2	-	9.52	10.52	-
1.1	-	-	-	-
1.0	-	-	-	-
0.9	-	-	5.26	-
0.75	-	-	-	-
0.20	6.66	-	-	-

Furthermore, the AS concentration resulting in the crystallization of four classes of proteins as complex structures was studied as only limited data for unique complex structures is available. The dataset includes only the protein complex entries having 30% sequence identity to accommodate the maximum available pool of distinct protein sequences/structures. The analysis indicates that the AS conc. leading to crystallization of proteins as complex structures showed a preferential pattern. The three protein classes as complex structures show the maximum crystallization percentage at 2.0M and a preferential pattern at other AS concentrations [Table 2]. 'All Alpha' proteins class as complex structures show a distinct preference at 1.7M, 2.2M & 2.5M leading to crystallization of appox. 40% of 'All Alpha' proteins as complex structures. 'All Beta' proteins class as complex structure show a distinct preference for only four AS conc. in an order of 2.0M>1.8M>1.2M=3.5M. It indicates that nearly 50% of 'All Beta' proteins crystallized as complex at only narrow range of two AS conc. i.e. 2.0M (33.33%) & 1.8M (19.04%). 'Alpha and Beta (α+β)' proteins class as complex structures shows preference in an order of 2.0M>1.5M>1.8M>2.8M. 'Alpha and Beta (α/β)' proteins class as complex structures shows preference in an order of 2.0M>1.5M=1.2M>1.6=1.9M>1.3M=1.4M=2.1M=2.4M=0.9M. 'Alpha and Beta (α+β & α/β)' protein classes show maximum crystallization as complex at three AS concentration. Alpha and Beta (α+β)' protein class shows 57.12% of crystallization as complex at three AS conc. i.e. 2.0M (23.80%)>1.5M (19.04%)>1.8M (14.28%). Alpha and Beta (α/β)' protein class shows 42.09% of crystallization as complex at three AS conc. i.e. 2.0M (21.05%)>1.5M (10.52%) =1.2M (10.52%) class of proteins. These four classes of proteins show 36-47% of crystallization percentage at three AS conc. i.e. 2.0M, 1.8M and 1.6M. 'Alpha and Beta (α/β & α+β)' protein class show similar percentage of crystallization i.e. ~42%, while 'All Alpha (47.81%)' & 'All Beta (36.82%)' proteins remains at two extreme at these three AS conc. The preferential pattern observed for the four protein classes as complex structures is influenced by type and chemical nature of the lig and/substrate/inhibitor/protein etc. and the resulting interplay of attractive and repulsive forces due to exposure of acidic and/or basic protein surface residues. Inspite of a preferential pattern for AS conc. is observed, broadly it can be concluded that the optimum AS conc. leading to crystallization of all the four classes of proteins as complex structures lie within a narrow range of 1.5M-2.5M as also observed above in case of single &soluble proteins. Therefore, this narrow AS conc. range can be utilized in the preparation of commercial screens having improved crystallization efficiency with lower protein requirement.

The two theoretical protein parameters i.e. iso-electric point (pI) and aliphatic index studied for the four classes of single & soluble proteins in order to understand the correlation between these protein parameters and AS conc. facilitated the crystallization of proteins. This study is not followed for complex structure as these are protein sequence based parameters. These two protein parameters are considered as there is an increased chance for a protein to crystallize near the pI of the protein moiety [15] and Aliphatic index is an indicator of thermo stability of a protein [16] and also used to predict the interaction with other molecules or surfaces, which might influence the AS conc. based crystallization of proteins. The results shows an inverse relation between iso-electric Point (pI) and AS conc. for 'All Alpha' and 'All Beta' proteins and direct correlation for 'Alpha and Beta' Proteins [Fig. 3, 4, 5, 6]. In case of 'All Alpha' and 'All Beta' proteins, the protein crystallizes at lower AS conc. with increase of iso-electric point of a protein. The overall slope of the two curves does not show a sudden change, though there is a noticeable difference

between the slope of the curve for 'All Alpha & All Beta' proteins. 'Alpha and Beta' proteins correlation curve shows that with increase of iso-electric point there is an increase of AS conc. requirement for crystallization of these classes of proteins. The slope of the curve is more acute in case of α/β proteins in comparison to α+β proteins. These results are in contrast as reported earlier for ovalbumin (an alpha &beta protein). These contrasting results are possibly due to the fact that in ovalbumin publication, there is an inverse correlation between ionic strength of a buffer and iso-electric point of a protein [17].

The curves between aliphatic index of a protein and AS conc. shows that the 'All Alpha' and 'Alpha and Beta' proteins possess inverse relationship, while 'All Beta' proteins possess direct correlation between the two parameters [Fig. 7, 8, 9, 10]. The slope of the curve is more acute in case of 'Alpha and Beta Proteins'.

The smooth steepness of the curve observed for AS conc. facilitated crystallization of four classes of proteins and iso-electric point/Aliphatic index of the proteins is due to the narrow range of AS conc. leading to protein crystallization.
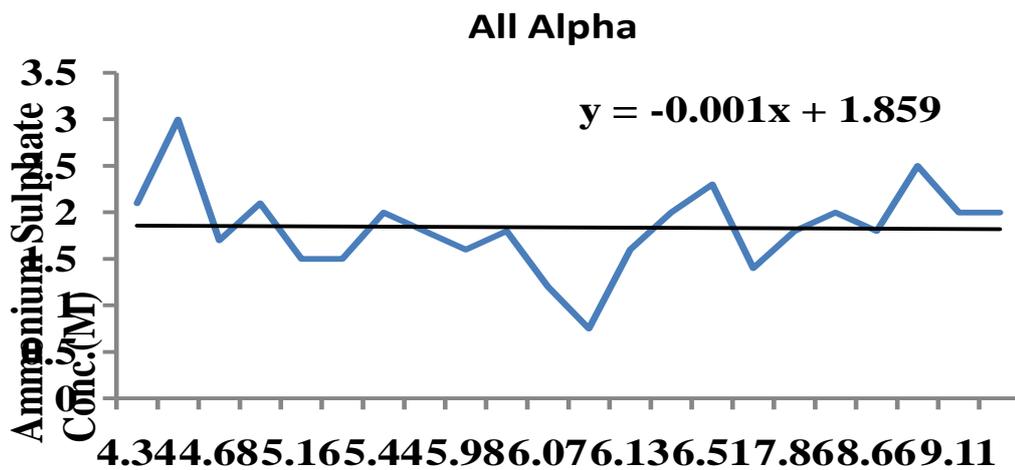
All Alpha

$$y = -0.001x + 1.859$$

Fig. 3: Shows the relation between Iso-electric point (pI) and Ammonium Sulphate (AS) Concentration (M) for Alpha Protein type. The Graph also shows the trend line equation.
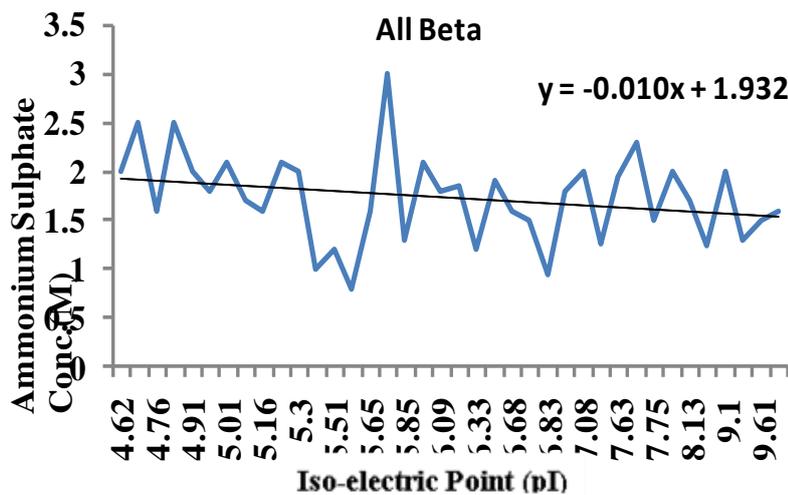..

All Beta

$$y = -0.010x + 1.932$$

Iso-electric Point (pI)

Fig. 4: Shows the relation between Iso-electric point (pI) and Ammonium Sulphate (AS) Concentration (M) for Beta Protein type. The Graph also shows the trend line equation.
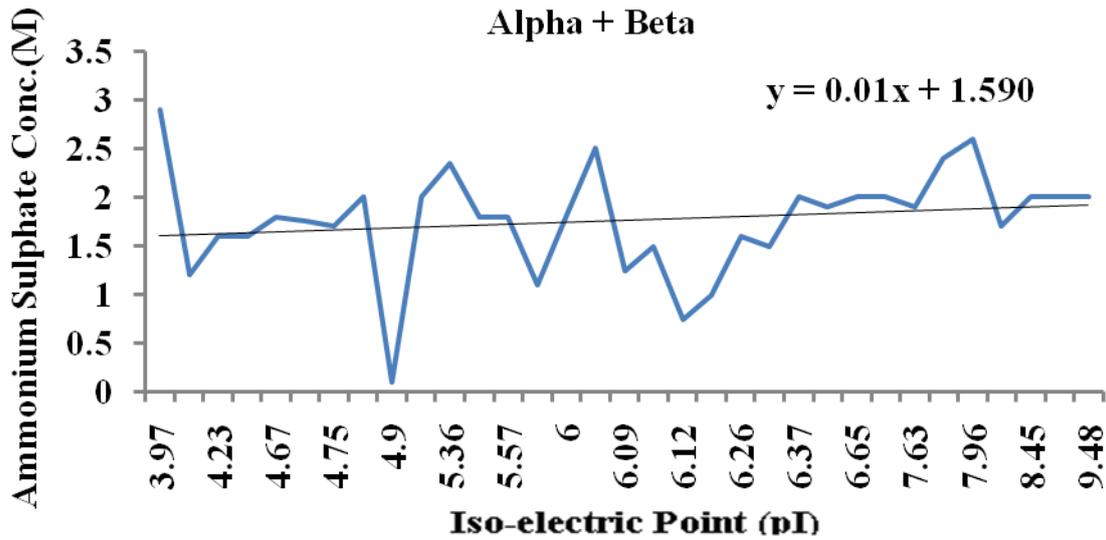..

32

Fig. 5: Shows the relation between Iso-electric point (pI) and Ammonium Sulphate (AS) Concentration (M) for Alpha and Beta (α+β) Protein type. The Graph also shows the trendline equation.
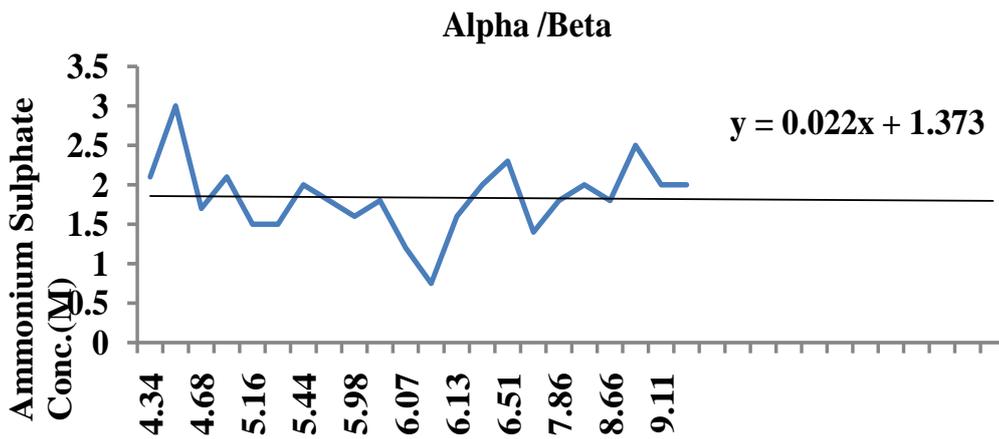
....

Fig. 6: Shows the relation between Iso-electric point (pI) and Ammonium Sulphate (AS) Concentration (M) for Alpha and Beta (α/β) Protein type. The Graph also shows the trendline equation.
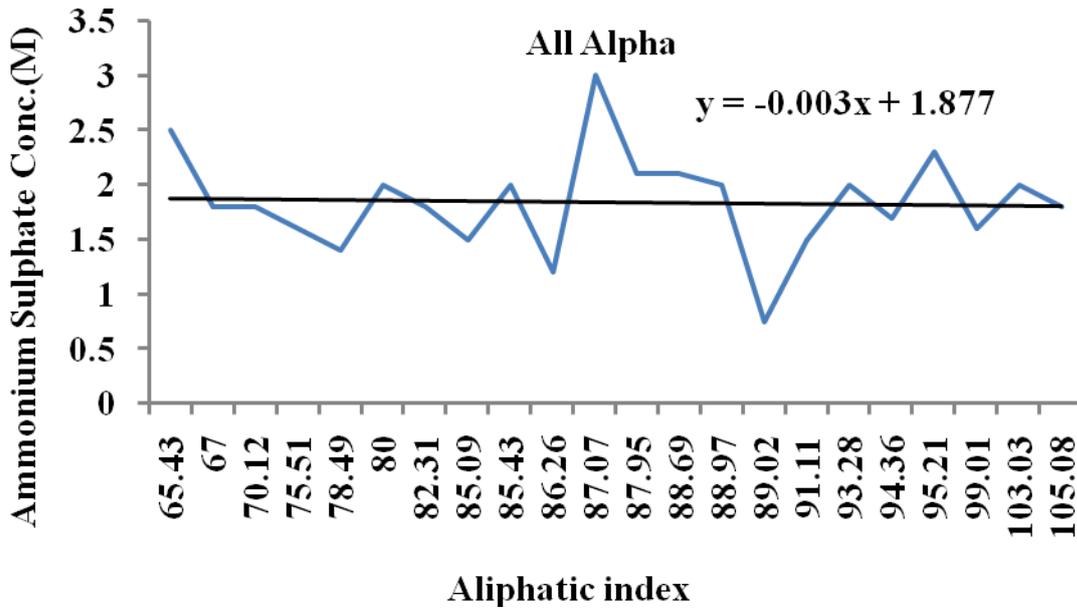
....

Fig. 7: Shows the relation between Aliphatic Index and Ammonium Sulphate (AS) Concentration (M) for Alpha Protein type. The Graph also shows the Trendline equation.

....
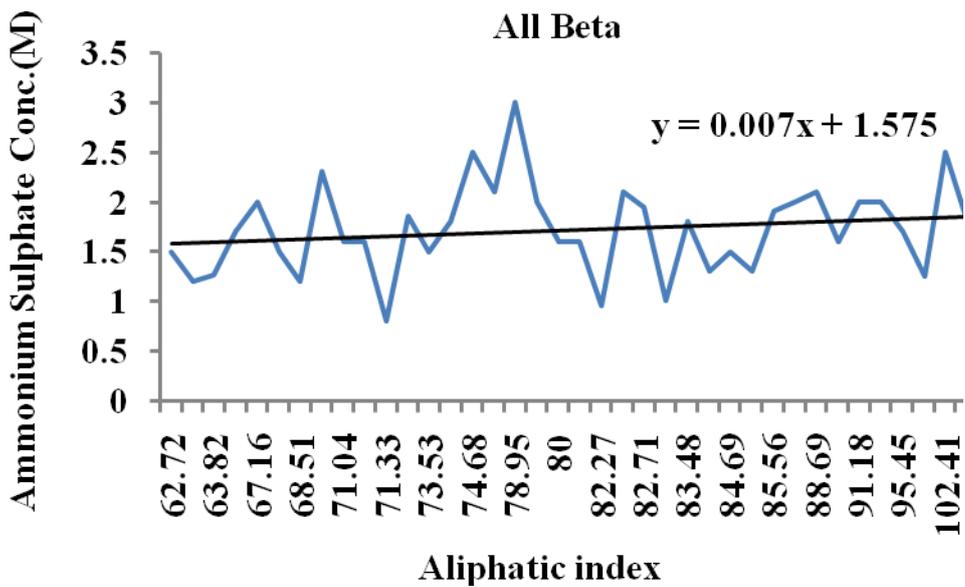

Fig. 8: Shows the relation between Aliphatic Index and Ammonium Sulphate (AS) Concentration (M) for Beta Protein type. The Graph also shows the trendline equation.
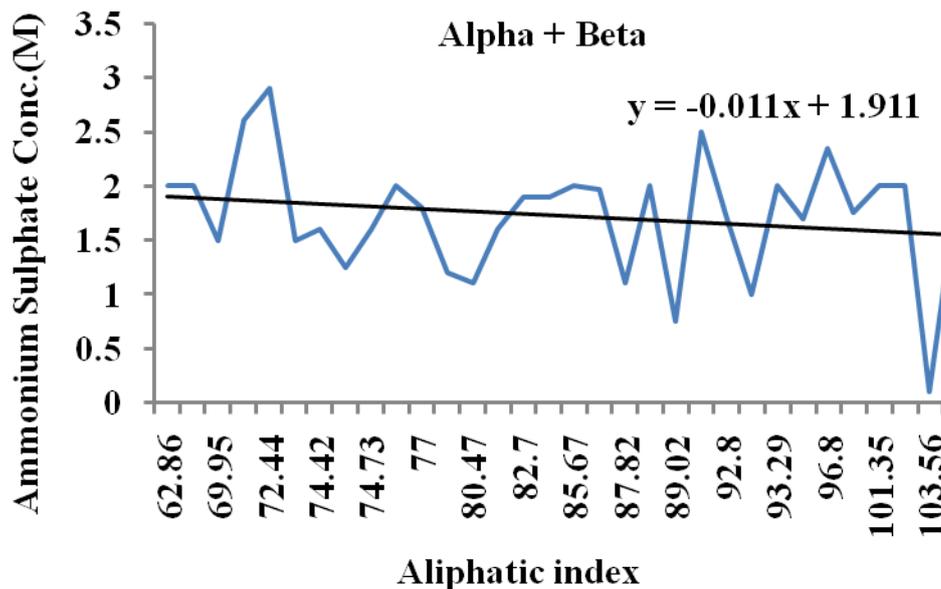
....

Fig. 9: Shows the relation between Aliphatic Index and Ammonium Sulphate (AS) Concentration (M) for Alpha and Beta (α+β) Protein type. The Graph also shows the trendline equation.

....
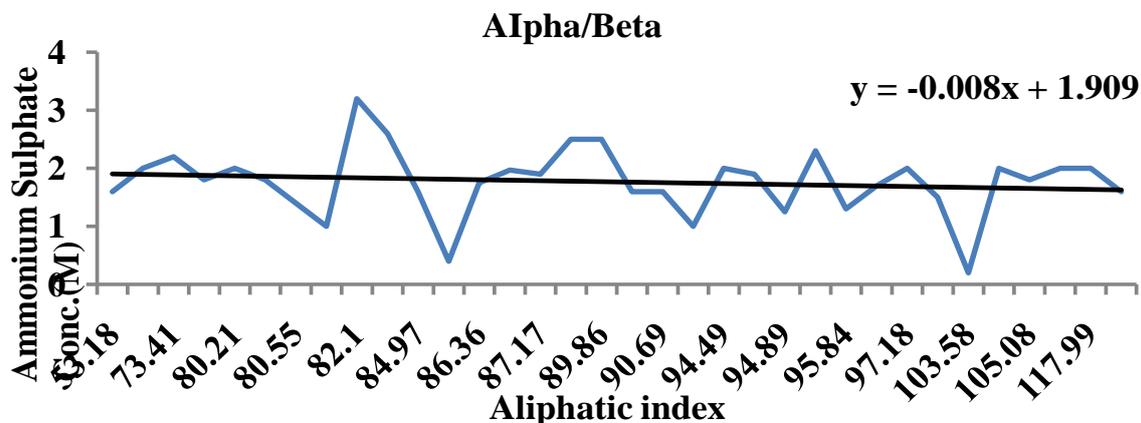

Fig. 10: Shows the relation between Aliphatic Index and Ammonium Sulphate (AS) Concentration (M) for Alpha and Beta (α/β) Protein type. The Graph also shows the trendline equation.

....

CONCLUSION

A number of AS based commercial screens are available and there is a scope of improving the efficiency of these screens. These results indicate that substantial percentage of four classes of proteins is crystallized within a narrow range of AS conc. i.e. 1.5-2.5M. Further the curves deduced between the two theoretical protein parameters i.e. isoelectric point and Aliphatic index and AS conc. suggest that these curves may be used as reference curves for determining the AS conc., which may facilitate protein crystallization for a particular class. Indeed, these results need empirical validation for improving the efficiency of crystallization process.

CONFLICT OF INTEREST
The author declares having no competing interest.

ABBREVIATIONS: PEG –Polyethylene Glycol, AS – Ammonium Sulphate

REFERENCES

[1] McPherson A. [1999] Crystallization of Biological Macromolecules. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[2] 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.

[3] 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.

[4] Newman J, Egan D, Walter TS, et al. [2005] Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy, Acta Cryst D. 61: 1426-1431.

[5] Luft JR, Newman J, Snell EH. [2014] Crystallization screening: the influence of history on current practice, ActaCryst F Struct Biol Comm. 70(7):835-853.

[6] Pereira JH, McAndrew RP, Tomaleri GP, Adams PD. [2017] Berkeley Screen: a set of 96 solutions for general macromolecular crystallization. J Appl Cryst. 50(5):1352-1358.

[7] Gaur RK. [2016] Estimation of PEG types and their concentration during protein crystallization. The IIOAB J 7(7):5-9

[8] Berman HM, Westbrook J, Feng Z, et al. [2000] The protein Data Bank Nucleic Acids Res 28:235–242.

[9] 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

[10] Gasteiger E, Hoogland C, et al. [2005] Protein Identification and Analysis Tools on the ExPASy Server; John M Walker (ed): The Proteomics Protocols Handbook, Humana Press 571-607.

[11] McPherson A. [1990] Current approaches to macromolecular crystallization. Eur J Biochem 189:1-23.

[12] Minor Dl Jr, Kim PS. [1994] Context is a major determinant of Beta sheet propensity, Nature 371(6494):264-267.

[13] Sitbon E, Pietrokovski S. [2007] Occurrence of Protein Structure Elements in conserved sequence regions. BMC Struct Biol 7:3.

[14] Dumetz AC, Sneelinger O'Brien AM, Kaler EW, Lenhoff AM. [2007] Patterns of Protein-Protein interactions in Salt solutions and implication for Protein crystallization, Prot Sci 16:1867-1877.

[15] Kantardjieff KA, Rupp B. [2004] Protein isoelectric point as a predictor for increased crystallization screening efficiency Bioinf 20:2162-2168.

[16] Ikai A. [1980]. Thermo stability and aliphatic index of globular proteins, J Biochem. 88:1895-1898.

[17] Smith ERB. [1935] The effect of variations in ionic strength on the apparent isoelectric point of egg albumin, J Biol Chem. 108:187-194.