Acta Chemica Scandinavica, 1996: 50: 68–70
Printed in Belgium – all rights reserved

ACTA
CHEMICA SCANDINAVICA
ISSN 0904–213X

The Solubility of Proteins in Organic Solvents

G. Houen

Statens Seruminstitut, Department of Autoimmunology, Artillerivej 5, DK-2300 S, Copenhagen, Denmark

Houen, G., 1996. The Solubility of Proteins in Organic Solvents. – Acta Chem. Scand. 50: 68–70 © Acta Chemica Scandinavica 1996.

The ability of organic solvents to solubilize proteins has been investigated: for bovine serum albumin, trifluoroacetic acid = formic acid > glycerol > 3-mercaptopropionic acid > dimethyl sulfoxide > triethanolamine > chloroform; for lysozyme, trifluoroacetic acid = formic acid = glycerol = dimethyl sulfoxide > 3-mercaptopropionic acid > triethanolamine > mercaptoethanol > DMF > acetic acid > chloroform > isoamyl alcohol > butyl chloride > 2-propanol.

The solubility of lysozyme and bovine serum albumin (BSA) in various organic solvents has been investigated. The solubility of lysozyme was found to be very high in the strongly polar, protic solvents formic acid and trifluoroacetic acid as well as in glycerol and dimethyl sulfoxide, but moderate to low in most dipolar aprotic solvents and very low in most apolar aprotic solvents. The strongly polar, protic solvents formic acid and trifluoroacetic acid were good at solubilizing BSA; glycerol was also a good solvent and dimethyl sulfoxide was moderately good at solubilizing BSA. Thus the size of the protein as well as the nature of the solvent are important factors for determining the solubility of proteins in organic solvents. These results are useful in the design of pharmaceutical protein formulations and in the design of procedures for chemical modifications and cleavage of proteins, and for extracting and manipulating proteins adsorbed to solid surfaces.

The folding of a polypeptide chain into the three-dimensional structure is an energetically favorable process in water, and most proteins have stabilities in the range 20–60 kJ mol^{-1}.[1] All information necessary to define the three-dimensional structure seems to be contained in the linear sequence of a protein, and thus the folding is specified by a 'folding code'.[2,3] In principle it should therefore be possible to predict the tertiary structure by knowledge of the primary structure. The folding and stability of a protein is, however, also dependent on the composition of the environment, and stability is defined as the difference in free energy of the denatured and native state, $G(D) - G(N)$.[4]

Among the agents known to affect protein stability are:[5] 1, detergents such as sodium dodecyl sulfate (SDS), which denatures many soluble proteins and alters the structure of some membrane proteins from, e.g., predominantly β-sheet to predominantly α-helix;[6] 2, chaeotropic agents such as urea, guanidine and LiCl; 3, organic solvents such as ethanol and methanol. Among the organic solvents, water-miscible solvents are most denaturing, but they also tend to precipitate proteins from solution. In contrast with this, many organic solvents that are not miscible with water, such as ethyl acetate and octane, have been found to be excellent media for many enzymatic reactions.[7]

In order to gain more insight into the relationship between solvent structure and protein stability and solubility, I have undertaken a study of the solubility of proteins in organic solvents. This knowledge is also useful for practical applications such as: 1, solid-phase peptide synthesis on proteins, recently introduced by us,[8] where new peptide chains are synthesized stepwise on protein lysine side chains, and where knowledge of the solubility of proteins in organic solvents is necessary; 2, protein sequencing, where prevention of washout of protein is essential during washing and reaction cycles; 3, recovery of proteins from membranes after electrotransfer from electrophoresis gels; 4, chemical modifications and cleavage of proteins in solution or on solid surfaces.

Materials and methods

Chemicals. Ethanol was from *De Danske Spritfabrikker* (Copenhagen, Denmark). Acetic acid, benzene, butanol, chloroform (stabilized with 1% ethanol), *N,N*-dimethylformamide (DMF), dimethyl sulfoxide (DMSO), dioxane, formic acid, glycerol, heptane, isoamyl alcohol, 2-propanol, methanol and 1-propanol were from Merck (Darmstadt, Germany). Butyl chloride and trifluoroacetic acid (TFA) were from Applied Biosystems (Foster City,

Table 1. Solubility of BSA in various organic solvents.

Solvent	Solubility/mg ml^{-1}	Solubility/μM	Comments
Acetic acid	0.02	0.29	clear supernatant
Acetonitrile	0.02	0.29	clear supernatant
Benzene	0.04	0.59	clear supernatant
Butanol	0.03	0.44	clear supernatant
Butyl chloride	0.01	0.15	clear supernatant
Chloroform (1% EtOH)	0.12	1.76	cloudy supernatant
N,-Dimethylformamide	0.01	0.15	clear supernatant
Dimethyl sulfoxide	5.12	75.29	clear supernatant
Dioxane	0.01	0.15	clear supernatant
Ethanol	0.01	0.15	clear supernatant
Formic acid	>50	>758	all dissolved
Glycerol	20.0	294.12	clear sup., half dissolved
Heptane	0.00	0.00	clear supernatant
Isoamyl alcohol	0.02	0.29	clear supernatant
2-Propanol	—	—	not determined
2-Mercaptoethanol	—	—	not determined
3-Mercaptopropionic acid	5.45	80.15	clear sup., some dissolved
Methanol	0.01	0.15	clear supernatant
1-Propanol	0.03	0.44	clear supernatant
Triethanolamine	0.70	10.29	clear supernatant
Trifluoroacetic acid	>50	>758	all dissolved

USA). Bovine serum albumin (BSA, 96–99% purity), lysozyme (95% purity), 2-mercaptoethanol, 3-mercaptopropionic acid and triethanolamine were from Sigma (St. Louis, USA). The purity of all solvents was of the highest commercially available grade.

Amino acid analysis. This was carried out according to Barkholt and Jensen (1989).[9]

Solubility experiments. Lysozyme (100 mg) or BSA (50 mg) was added to 1 ml solvent and the mixture was incubated at room temperature with 'end over end' agitation. After 24 h the solution was centrifuged (1000 *g*, 15 min) and 10 or 100 μl aliquots were withdrawn in duplicate, dried by evaporation of the solvent in a nitrogen stream, and subjected to hydrolysis in 6 M HCl and subsequent quantitative amino acid analysis. The amount of protein in the samples was then calculated from the leucine content of the hydrolysates and from the known M_r and amino acid composition of the proteins used.

Results

Previous experiments aimed at solubilizing denatured proteins, and recovering polypeptides from membranes, indicated that only highly polar solvents in combination with water could solubilize the proteins.[10] To investigate

Table 2. Solubility of lysozyme in various organic solvents.

Solvent	Solubility/mg ml^{-1}	Solubility/μM	Comments
Acetic acid	0.55	38.46	clear supernatant
Acetonitrile	0.03	2.10	clear supernatant
Benzene	0.02	1.40	clear supernatant
Butanol	0.03	2.10	clear supernatant
Butyl chloride	0.08	5.59	clear supernatant
Chloroform (1% EtOH)	0.25	17.48	clear sup., some dissolved
N,-Dimethylformamide	0.61	42.66	clear supernatant
Dimethyl sulfoxide	>100	>7000	clear sup., all dissolved
Dioxane	0.02	1.40	clear supernatant
Ethanol	0.02	1;40	clear supernatant
Formic acid	>100	>7000	all dissolved
Glycerol	>100	>7000	all dissolved
Heptane	0.02	1.40	clear supernatant
Isoamyl alcohol	0.08	5.59	clear supernatant
2-Propanol	0.07	4.90	clear supernatant
2-Mercaptoethanol	0.65	45.45	clear sup., some dissolved
3-Mercaptopropionic acid	4.20	293.71	clear sup., much dissolved
Methanol	0.01	0.70	clear supernatant
1-Propanol	0.02	1.40	clear supernatant
Triethanolamine	0.88	55.94	clear supernatant
Trifluoroacetic acid	>100	>7000	all dissolved

the solubility in more detail bovine serum albumin (BSA, M_r 66000) was chosen as a representative of a large multi-domain protein, and lysozyme (M_r 14300) was chosen as a representative of a small single-domain protein.

Table 1 shows the results of determining the solubility of BSA in organic solvents. As seen the solubility in most solvents is practically zero, including the solvents most used for reversed-phase chromatography of proteins: CH_3CN and MeOH. However, a modest but significant solubility was seen for chloroform (stabilized with 1% ethanol) and a somewhat higher solubility was seen for triethanolamine. Among solvents with a good to high solubilizing power the order of effectiveness was: trifluoroacetic acid = formic acid > glycerol > 3-mercaptopropionic acid > dimethyl sulfoxide > triethanolamine > chloroform.

For lysozyme essentially the same pattern was observed (Table 2). However, a low, but significant, solubility was observed in DMF, CH_3COOH and mercaptoethanol in addition to chloroform and triethanolamine. Those solvents that are good solvents for BSA were also found to be good solvents for lysozyme. The order of effectiveness for solvents solubilizing lysozyme was found to be: trifluoroacetic acid = formic acid = glycerol = dimethyl sulfoxide > 3-mercaptopropionic acid > triethanolamine > mercaptoethanol > DMF > acetic acid > chloroform > isoamyl alcohol > butyl chloride > 2-propanol. The solubility of lysozyme was substantially higher than that of BSA and some solvents which solubilized lysozyme hardly solubilized BSA at all (mercaptoethanol, DMF, acetic acid).

Discussion

Solvents can be grouped into four categories: polar protic solvents, dipolar aprotic solvents, apolar protic solvents and apolar aprotic solvents.[11] Among these categories the polar protic solvents have until recently been regarded as the only solvents capable of dissolving proteins.[12–14] This dogma has been reinvestigated in parallel with the present work by Chin *et al.*[13] and Bromberg and Klibanov,[14] who found that some proteins could be solubilized in apolar protic solvents provided that the proteins were lyophilized at a pH far from the isoelectric point prior to solvation. The reason for this observation may be a purely electrostatic, aggregation diminishing effect, or may be due to a larger amount of water retained by the lyophilized protein. I have chosen to compare two commonly used proteins as representatives of a large multi-domain protein (BSA) and a small single-domain protein (lysozyme). The proteins were used as the commercial preparations which give a pH of 7 when dissolved in water. In this way, the results are directly comparable with the 'natural' solvent, water. The polar protic solvents were superior to other solvents in dissolving BSA and lysozyme, and TFA and formic acid were the only solvents capable of completely dissolving the amount of BSA used (50 mg/ml). Somewhat surprisingly, acetic acid was almost ineffective in dissolving BSA and lysozyme while 3-mercaptopropionic acid only had a moderate solubilizing power for BSA. These solubilizing abilities are most readily explained by differences in dielectric constants and hydrogen bonding properties of the solvents. Good solvents belong to the same class as water and comprise TFA, HCOOH and glycerol. Thus, the high solubility of both proteins in glycerol can most readily be ascribed to the 'water-like' structure of this solvent. Among the dipolar, aprotic solvents DMSO has the highest solubilizing power, probably due to its high dielectric constant ($\varepsilon = 46.7$), while DMF was only a moderately good solvent ($\varepsilon = 36.7$).

Altogether, the results point to the importance of the hydrogen-donating ability of the solvent for solubilizing proteins. Do these results have any relevance for understanding protein stability? It can be assumed that proteins have evolved under the demand to be both stable and soluble in water. This is achieved by exposing protic groups on the surface and by shielding apolar aprotic groups in the interior of the protein. It would then seem that stability and solubility are complementary properties in polar protic solvents, and this would imply that proteins should retain their native structure in TFA, formic acid and glycerol, and possibly also in the dipolar, aprotic solvent DMSO. With regard to protein formulations for pharmaceutical use glycerol would seem to be a solvent that deserves further attention. Disregarding its high viscosity, this solvent has many desirable properties for biological applications as it is non-toxic and is known to stabilize many proteins.[5]

References

1. Pace, C. N. *Trends Biochem. Sci. 14* (1990) 14.
2. Rossmann, M. G. and Argos, P. *Ann. Rev. Biochem. 50* (1981) 497.
3. Jaenicke, R. *Prog. Biophys. Mol. Biol. 49* (1987) 117.
4. Privalov, P. L. and Gill, S. J. *Adv. Protein Chem. 39* (1988) 191.
5. Volkin, D. B. and Klibanov, A. In: Rickwood, D. and Hames, B. D., Eds., *Protein Function, a Practical Approach*, IRL Press, Oxford, UK 1989, pp. 1–24.
6. Eisele, J.-L. and Rosenbusch, J. P. *J. Biol. Chem. 265* (1990) 10217.
7. Klibanov, A. M. *Trends Biochem. Sci. 15* (1989) 141.
8. Hansen, P. R., Holm, A. and Houen, G. *Int. J. Peptide Protein Res. 41* (1993) 237.
9. Barkholt, V. and Jensen, A. L. *Anal. Biochem. 177* (1989) 318.
10. Houen, G. *Methods Mol. Cell Biol. 4* (1993) 105.
11. Lowry, T. H. and Richardson, K. S. *Mechanism and Theory in Organic Chemistry*, Harper and Row, New York, USA 1976.
12. Singer, S. J. *Adv. Protein Chem. 17* (1992) 1.
13. Chin, J. T., Wheeler, S. L. and Klibanov, A. M. *Biotechnol. Bioeng. 44* (1994) 140.
14. Bromberg, L. E. and Klibanov, A. M. *Proc. Natl. Acad. Sci. USA 92* (1995) 1262.

Received May 17, 1995.