The Journal of Engineering Research (TJER) Vol. 12, No. 2 (2015) 01-14

Arginine-Amino Acid Interactions and Implications to Protein Solubility and Aggregation

A. R. Shaikh*[a] and D. Shah[b]

[a] Department of Chemical Sciences and Engineering, Kobe University, Japan
[b] Petroleum and Chemical Engineering Department, Sultan Qaboos University, Oman

Received 23 February 2014; Accepted 23 February 2015

Abstract: Arginine, useful in protein refolding, solubilization of proteins, and suppression of protein aggregation and non-specific adsorption during formulation and purification, is a ubiquitous additive in the biotechnology and pharmaceutical industries. In order to provide a framework for analyzing the molecular level mechanisms behind arginine/protein interactions in the above context, density functional theory was used to systematically examine how arginine interacts with naturally occurring amino acids. The results show that the most favorable interaction of arginine is with acidic amino acids and arises from charge interactions and hydrogen-bond interactions. Arginine is also shown to form stacking and T-shaped structures with aromatic amino acids, the types of cation–π and N–H···π interactions, respectively, known to be important contributors to protein stability. The analysis also shows that arginine-arginine interactions lead to stable clusters, with the stability of the clusters arising from the stacking of the guanidinium part of arginine. The results show that the unique ability of arginine to form clusters with itself makes it an effective aggregation suppressant and support the interpretations of the current study using experimental and molecular dynamics results available in the literature. The results also contribute to understanding the role of arginine in increasing protein solubility, imparting thermal stability of important enzymes, and designing better additives.

Keywords: Protein aggregation, Protein solubility, Arginine, Quantum chemical calculation, Amino acid.

تفاعلات الأرجينين مع الأحماض الأمينية وردود فعلها على تكدس وذوبان البروتين

عبدالرزاق شيخ، ودهول شاه

المستخلص: يعد الأرجينين، المفيد في طي وذوبان البروتينات، وإخماد تكدسها، والامتزاز غير المحدد لها أثناء تركيبها وتنقيتها، من مواد الإضافة المحسنة الواسعة الانتشار في صناعات التكنولوجيا الحيوية والدوائية. في هذه الدراسة، ومن أجل توفير إطار لتحليل آليات المستوى الجزيئية وراء تفاعلات الأرجينين / البروتين في السياقات السالفة الذكر، تم استخدام نظرية كثافة الدالة من أجل الفحص بمنهجية لكيفية تفاعل الأرجينين التي تحدث مع الأحماض الأمينية المتكونة تلقائيًا بحكم الطبيعة. لقد أظهرت النتائج أن التفاعل الأكثر ملاءمة مع الأرجينين هو مع الأحماض الأمينية الحمضية وينشأ ذلك بسبب الشحنات الكهربية والروابط الهيدروجينية. ولقد أظهر كذلك الأرجينين مع الأحماض الأمينية العطرية تفاعلات من نوع التراص والتراكيب على شكل حرف ت، ومن المعروف أن التفاعلات من نوع كاتيون- باي ونوع ن- اتش. باي لها مساهمات مهمة في استقرارية البروتين. ويظهر التحليل أيضا أن تفاعلات الأرجينين- الأرجينين تؤدي إلى مجموعات عنقودية مستقرة، وإن هذه المجموعات العنقودية المستقرة ناشئة عن التراص من جهة الجوانيدين والذي هو جزء من الأرجينين. ولقد استخدمنا نتائج هذه الدراسة لعرض قدرة الأرجينين الفريدة في تشكيل مجموعات عنقودية مع نفسها مما يجعل منه مانعا فعالاً لتكدس البروتينات وقد دعمنا هذا التفسير بنتائج من التجارب ومن محاكاة الديناميكا الجزيئية المتاحة في الأدبيات (المراجع). وفي هذا البحث علقنا كذلك على الملابسات واستخدامنا النتائج لفهم دور الأرجينين في زيادة ذوبان البروتين ومنح الاستقرار الحراري للإنزيمات الهامة ذات الشأن، وكذلك في تصميم مضافات أفضل.

الكلمات المفتاحية: تكدس البروتين، ذوبان البروتين، الأرجينين، حسابات الكم الكيميائية، الأحماض الأمينية.

Corresponding author's e-mail: abdulrajjak.shaikh@people.kobe-u.ac.jp

1. Introduction

Amino acids are critical to life as they play a central role in biology and physiology, both as building blocks of proteins and as intermediaries or vehicles in various metabolisms. Proteins catalyze almost all intracellular reactions and also control most of the cellular processes. They are active only in their respective, unique three-dimensional (3D) structures. Interactions between amino acids, contained within the protein sequence, carry the necessary information to fold the proteins correctly and stabilize the active structure. For example, a point mutation in the amino acid sequence within a protein can substantially change the interactions and thus the structure and functionality of the protein molecules (Leippert *et al*. 1997). Therefore, there is considerable interest in a systematic examination of amino-acid/amino-acid interactions in order to understand the structural stability of proteins.

The problem of protein *aggregation* is intimately tied to protein folding and the stability of the folded shape and is important not only in cell biology but also in pharmaceutical, recombinant proteins, and the food industry. During the past two decades, the advent of recombinantly expressed therapeutic proteins in the pharmaceutical industry has highlighted the importance of issues such as the chemical and physical stability of the resulting proteins and their stability during long-term storage (Frokjaer and Otzen 2005). Proteins tend to degrade through aggregation with even a slight change in environmental conditions, including temperature, pH, buffers, and ion concentration. An increasing number of disorders, including Alzheimer's and Parkinson's disease, the spongiform encephalopathies, and type two diabetes, also are directly associated with the deposition of protein aggregates in tissues, including the brain, heart and spleen (Cleland 1991; Cellamare *et al.* 2008).

A commonly used technique to prevent protein misfolding and aggregation is the use of low-molecular-weight additives, termed "artificial molecular-chaperones". Artificial molecular chaperones have been developed to prevent aggregation *in vitro* and promote efficient refolding of denatured proteins via the addition of chemical compounds, including amino acids (arginine, proline, alanine, etc.), denaturants (typically guanidine and urea), poly(ethylene glycol), surfactants, and so on (Chen *et al.* 2009; Lee *et al.* 2006; Ghosh *et al.* 2009; Matsuoka *et al.* 2007; Shiraki *et al*. 2002; Takano *et al.* 2009). Clearly, the intermolecular interactions between these stabilizers and the amino acids on the protein surface play a dominant role in stabilizing the proteins against aggregation and, therefore, deserve careful attention so that educated choices for stabilizers could be made for a given protein.

One of the most commonly used additives for protein stabilization is arginine, which is effective in suppressing protein aggregation, increasing the refolding yield of proteins, and enhancing the solubility of aggregation-prone proteins as required for pharmaceutical applications (Shiraki *et al.* 2002). Apart from this, arginine also finds a number of applications in improving the yield of ion and affinity chromatography (Ejima *et al.* 2005). Arginine is also used to protect enzymes from thermal denaturation. An example is phytase, an enzyme which is added to animal feed and then subjected to thermal processing in which the feed is heated to 60–90°C. However, as phytase is a heat-unstable enzyme which loses its activity during processing, arginine is added to the pre-feed formulation in order to prevent thermal denaturation. The addition of even as little as 1% arginine to the feed formulation has been shown to remarkably enhance the stability of enzyme (Ryu and Park 1998).

Although arginine is a frequent choice as an aggregation suppressor, refolding enhancer, and protein solubilizer, the exact stabilization mechanism of how arginine preserves enzyme stability is still unclear, although it is suspected that arginine binds to some regions of the protein via ionic interactions. A number of different hypotheses to explain the mechanism have been reported, including preferential interactions with the protein surface, increase of surface tension, and protein hydration (Arakawa *et al.* 2007). Different propositions recently have been advanced to explain the stabilizing effect of arginine. For example, a recent experimental study by Das *et al.* (2007) suggests that arginine stabilizes proteins against aggregation by reducing the hydrophobic interactions between the proteins through hydrophobic association between the aliphatic chain of arginine and hydrophobic patches on the protein, whereas Ghosh *et al.* (2009) suggest that strong interactions of arginine with tryptophan residues are the ones responsible for the stabilization. In general, the presently available studies do not explain or illustrate the mechanistic basis of arginine as a protein stabilizer adequately.

It is well known that the amino acid sequence of a particular protein contains all the information needed to fold protein in a specific conformation,

which ensures effective interaction within itself as well as with specific receptor sites required to trigger the biological response or functioning of the protein. Interactions with the aqueous solvent, known as the hydrophobic effect, result in residues with non-polar side chains typically being buried in the interior of a protein. Conversely, polar amino acid side chains tend to collect on the surface of a protein, where they are exposed to the aqueous medium. Hence, it was assumed that when therapeutic protein is in its native conformation, interaction between the side chain of exposed amino acids and the arginine will play an important role in stabilizing protein from aggregation. Considering this fact, only a truncated model of a side chain was used in the present calculations. Such a truncated model has been used successfully to study amino acid side chain interactions. Herewith, in order to develop a framework for subsequent analyses of the role of arginine in the above contexts (suppressing protein aggregation, preventing thermal denaturation, and enhancing protein solubility), the atomic level interactions between arginine and each amino acid are studied using density functional theory and a comparative picture of the interaction of arginine with each amino acid is presented. The present analysis sheds light on the different kinds of hydrogen bonds and the dominant interactions that exist between arginine and amino acids in a protein, some of which have been neglected in the past (for example, N–H$\cdots\pi$ interactions) so that one can identify which of the many possible interactions are either favorable or unfavorable for a particular application. Since, in general, favorable interactions lead to an increase in solubility of the protein and suppression of protein aggregation (Shiraki *et al.* 2002), and unfavorable interactions have the reverse effect (Timasheff 2002), the present study and similar studies for other additives can assist in designing protein-specific strategies to prevent the aggregation and misfolding of proteins, enhance protein solubility, and develop generally accepted guidelines and tests so that unforeseen adverse effects can be avoided by choosing additives most appropriate for a given protein.

2. Computational Methods

All the calculations presented here were carried out using the DMol3 numerical density functional theory program (Delley 1990; Delley 2000) computed in Materials Studio 4.3 (Accelrys Inc., San Diego, California, USA). The double numerical polarized (DNP) basis set, which includes all occupied atomic orbitals plus a second set of valence orbitals and polarized d-valence orbitals, was employed. Geometrical optimizations were carried out using the local density approximation (LDA) and the Vosko-Wilk-Nusair (VWN) exchange correlation functional. Optimized geometries were refined further by calculating single-point energy using the general gradient potential approximation (GGA) and the Becke-Lee-Yang-Parr (BLYP) exchange correlation functional. Charges and spin populations were calculated by the Hirshfeld and Mulliken population analysis. The effect of the aqueous environment was accounted approximately by computing single-point energies using the COnductor-like Screening MOdel (COSMO) (Klamt and Schuurmann 1993; Klamt 1995) at the GGA-Becke-Lee-Yang-Parr (BLYP) level. Frequency calculations were performed to ensure that structures obtained were indeed minimum energy structures with zero imaginary frequencies.

Charge transfer and chemical bonds are, in most cases, well described within design for testing (DFT) with either the local density approximation (LDA) or the GGA to account for the exchange correlation energy of the electrons. However, DFT does not include London-dispersion forces (Ortmann *et al.* 2005; Chakarova-Kack *et al.* 2006), which represent the dominant stabilization energy contribution required for stacking-type interactions. While it has recently become possible to include such effects in a non-empirical way in DFT calculations for systems of practical interest (Chakarova-Kack *et al.* 2006), such calculations are still computationally very expensive. However, dispersion interactions are compensated in DMol^3 by other terms in the interaction energy (Andzelm *et al.* 2003; Natsume *et al.* 2006). In *ab initio* calculations of interaction energies, the basis set superposition error (BSSE) approach is used to account for incomplete atomic basis sets. The BSSE contribution is expected to be small, as the DMol3 program uses numerical functions that are far more complete than the traditional Gaussian functions (Andzelm *et al.* 2003).

The configurational binding energy E_b between arginine and an amino acid was calculated by the following equation:

$$E_b = [Arg :: AA] - [Arg] - [AA] \quad (1)$$

where [Arg] stands for the energy of the arginine molecule, [AA] is the energy of the amino acid, and [Arg::AA] represents the energy of the arginine-amino acid complex. The interaction energy includes the effect of zero-point vibrational energy (ZPVE), which was obtained from normal mode analysis. Gibbs free energy at 298.15 K (room temperature) was calculated using the vibrational analysis from DMol[3].

The focus of the current study is on the interactions of the guanidinium group of arginine with the side-chains of amino acids with the arginine molecule truncated at the methyl guanidine part in order to focus particularly on the interactions of the guanidinium group of arginine with various protein residues and reduce the computation time and complexity arising from the flexibility of the side-chain of arginine. Further, focus has been restricted to the interactions of the methyl guanidine part of arginine with the side-chains of different residues, as guanidine is known to not interact with the protein backbone (Lim *et al.* 2009). Similar truncated systems have been successfully used previously to demonstrate some of the interactions between amino acids (Melo *et al.* 1999). Different possible orientations of each amino acid relative to arginine were studied in order to identify the maximum binding energy, both at 0°K and at room temperature. Because of the positive charge on arginine at neutral pH, the protonated form of arginine was used in the calculations; however, the neutral form was also considered wherever appropriate.

3. Results and Discussion

Section 3.1 begins with the results of the DFT calculations for the binding energies for arginine's interactions with itself and with the various amino acids. The strengths of binding and the most favorable orientations and structures are discussed. This is followed by brief discussions of other interactions (*ie*. beyond interaction of arginine with a single amino acid) such as interactions with more than one amino acid simultaneously in Section 3.2. Section 4 then focuses on a particular application, namely, implications of some observations of the role and mechanism of arginine as a protein stabilizer against aggregation, in protein solubility, and the thermal stability of enzymes.

Table 1. Classification of amino acids based on their side chains.

Group	Classification	Amino acids
1	Acidic and their amides	Asp, Glu, Asn, Gln
2	Ring containing	Trp, Phe, Tyr, Pro
3	Non-aromatic hydroxyl R-groups	Ser, Thr
4	Basic groups	Arg, Lys, His
5	Sulfur groups	Cys, Met
6	Aliphatic groups	Ala, Val, Leu, Ile

It is convenient to classify the amino acids into six different groups and calculate the binding energy of each amino acid with arginine. Group 1 consists of amino acids with acidic and amine-containing side chains (Asp, Glu, Asn, and Gln); Group 2 contains amino acids with rings (Phe, Tyr, Trp, and Pro); Group 3 is non-aromatic amino acids with hydroxyl R-groups (Ser and Thr); Group 4 is basic amino acids (Lys, Arg, and His); Group 5 is amino acids with sulfur-containing side chains (Cys and Met); and Group 6 contains amino acids with aliphatic R-group side chains (Gly, Ala, Val, Leu, and Ile) [Table 1].

3.1 Interaction of Arginine with Single Amino Acids

Group 1: Acidic amino acids and their amides

The interaction of the side-chains of amino acids in proteins is a determining factor of the mechanisms behind a wide variety of biological phenomena such as antigen/antibody recognition and enzyme-substrate interactions. In particular, the interactions involving ionic groups of opposite charges are expected to be more dominant because of the electrostatic contributions. One special case of these interactions involves the guanidinium group of the arginine and carboxylate group of acidic amino acids—forming a guanidinium-

carboxylate salt bridge. A detailed theoretical study of the interaction of arginine with acidic amino acids has been presented previously by Melo *et al.* (1999), who showed that the zwitterionic form of amino acids is more stable in an aqueous environment while the neutral form is more stable *in vacuo*. A similar situation was observed in the present calculations as well as strong interactions of arginine with acidic amino acids in the aqueous environment. The calculations of the current study indicate that preferred conformations of the arginine/amino acid pair are those where it is possible to establish two hydrogen bonds of the type N–H···O=C, where the two oxygen atoms of the carboxylate group of acidic amino acids share a partial negative charge and the two N–H groups of arginine carry a partial positive charge. There are two such interactions possible involving different nitrogen atoms of arginine. The results show a pronounced difference between the two conformations, with the hydrogen bonds of carboxylate with η_1N and η_2N being preferred over the hydrogen bonds with εN and η_2N of the arginine. The difference between the binding energies of the two conformations is 26.21 kcal/mol. Interactions between arginine and amino acids with amide side chains (Asn and Gln) are weaker than those with acidic amino acids (Figures 1A and B). The difference in binding energies between acidic and amino-side-chain-containing amino acids was found to be ~10 kcal/mol [Table 2], since only one hydrogen bond is possible in the latter case. In summary, arginine has a stronger interaction with acidic amino acids than with amido-side-chain-containing amino acids.

The results shown above are based on calculations done at ground state (0 °K). Additional calculations were performed to include the effects of entropy and temperature (298.15°K). The Gibbs free energies of binding ($G_b^{298.15K}$) of arginine with the various amino acids are also presented in Table 2, which shows that even at 298.15°K, the interactions between arginine and acidic amino acids dominate those with other amino acids.

Group 2: Amino acid-containing rings

Arginine forms parallel stacking interactions with aromatic amino acids (Flocco and Mowbray 1994). These types of aromatic-arginine interactions, which are often found in locations critical to the activities of proteins, apparently serve to orient the arginine side chain without interfering with its ability to form hydrogen bond elsewhere. Long-range interactions between arginine and aromatic residues have been reported by Martis *et al.* (2008); here, the cation from arginine interacts with the π-cloud of the aromatic rings, forming a stacking cation-π interaction. In this type of interaction, the side chains of aromatic amino acids provide a surface of negative electrostatic potential that can bind to a wide range of cations through predominantly electrostatic interactions. Martis *et al.* (2008) performed detailed analyses on a number of different proteins to show that 31% of the total Trp residues within the protein structure participate in such interactions with arginine. These types of cation–π interactions are more hydrophobic in nature than electrostatic. Martis *et al.* (2008) also observe that, amongst all the aromatic residues, the average energy of Arg-Trp interactions is the highest; however, Phe is shown to have the maximum number of such interactions. Recently, Ottiger *et al.* (2009) also studied hydrogen bonding interactions between amino and aromatic moieties. They suggested that the formation of such a T-shape interaction is due to the hydrogen bonding of a N–H···π nature. These interactions are amongst the weak intermolecular interactions found in proteins and play an important role in biological systems. Although Flocco and Mowbray (1994) suggest that parallel stacking interactions between arginine and aromatic residues are commonly found in protein structures and have high binding energies, the DFT calculations of the current study showed that T-shape interactions are more favored than the parallel stacking interactions by approximately 1.0 kcal/mol. This indicates that there is a competition between the parallel π stacking and N–H···π hydrogen bond (T-shape) interactions (Figure 1E and F). The highest E_b of arginine with Trp was observed in Group 2 amino acids [Table 2] because the indole group of Trp provides a larger negative electrostatic potential than benzene or phenol, thus making Trp constitute a more attractive cation-binding site.

Proline is a small molecule with a non-aromatic ring, and has therefore been included in the Group 2 classification. The binding energy of arginine with proline is –19.67 kcal/mol, higher than the one corresponding to tyrosine (–18.75 kcal/mol). The binding energy between Arg and Pro falls in between those of parallel stacking and T-shape interactions due to the flexible nature of the ring [Fig. 1H].

In summary, the binding energy of arginine with respect to amino acids with rings is in the following order [Table 2]:

Trp > Pro > Phe > Tyr

Ito *et al*. (2011), based on X-ray studies, reported that lysozyme mostly interacts with arginine. Shah *et al*. (2011; 2012) reported similar results using experimental and molecular dynamics studies.

Group 3: Amino acids with non-aromatic hydroxyl R-groups

The hydroxyl side chain of serine/threonine in a protein can establish additional intramolecular hydrogen bonds; in particular, it can act as a proton donor as well as a proton acceptor. Calculations from the current study show that the binding energy for arginine with serine is –22.70 kcal/mol and for threonine it is –21.11 kcal/mol. Strong binding energies were observed between arginine

Table 2. Binding energy (E_b at 0 K) and Gibbs free energy of binding (G_b at 298.15 K) of each amino acid with arginine calculated at GGA-PW91 level. All the energies are in kcal/mol.[a]

Group	Amino acid	E_b^{0K}	$G_b^{298.15K}$
1	Asp	–35.52	–20.84
	Glu	–34.72	–21.03
	Asn	–23.81	–12.51
	Gln	–24.39	–13.69
2	Trp	–17.82	–4.17
	Tyr	–17.50	–4.18
	Phe	–17.34	–4.50
	Pro	–18.01	–4.95
3	Ser	–19.83	–10.03
	Thr	–18.61	–5.28
4	Arg	–30.57	–20.84
	Lys	–12.54	–1.23
5	His	–23.11	–10.64
	Cys	–23.42	–10.10
	Met	–18.67	–5.14
6	Ala	–15.05	–6.13
	Val	–13.49	–0.69
	Leu	–14.31	–1.36
	Ile	–12.92	–0.04

Note: [a] Both the binding energy and the Gibbs free energy reported here are inclusive of ZPVE correction.

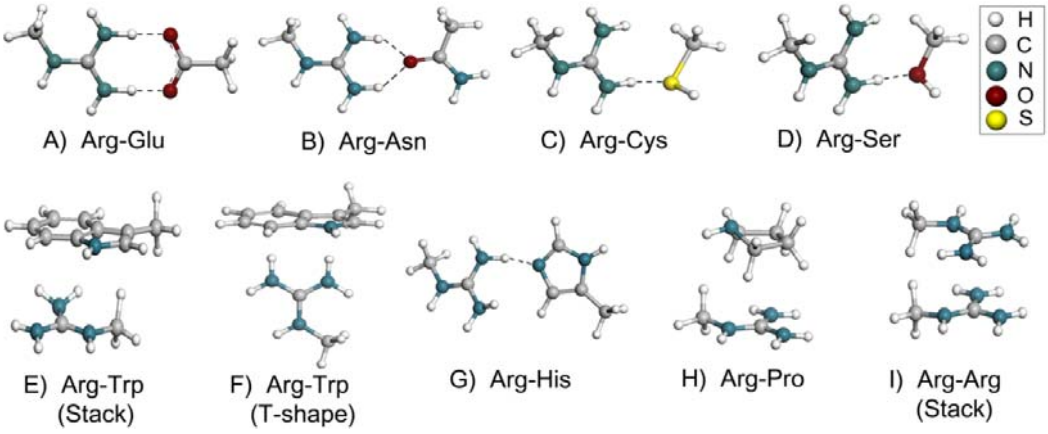

Figure 1. Optimized geometries of arginine-amino acid complexes. Only representative complexes are shown.

and Ser/Thr in a conformation where the hydrogen in the hydroxyl group of Ser/Thr points away from the arginine guanidinium group, hence making an N–H···O type of hydrogen bond [Fig. 1D]. Thus, electrostatic interactions and hydrogen bonding play a dominant role in the stabilization of such complexes. Also were calculated interactions of these amino acids with the neutral form of arginine; however, the binding energies of the resulting complexes are smaller (–9.20 kcal/mol for serine, –4.90 kcal/mol for threonine) than the ones corresponding to the cationic form of arginine. Hence, Ser/Thr also demonstrates a favorable interaction with arginine through N-H···O electrostatic interactions.

Group 4: Amino acids with basic groups

Although it is intuitive that arginine would have repulsive interactions with basic amino acids because both arginine and the basic amino acids have similar charges, a strong binding was observed in the form of parallel stacking alignment due to van der Waals interactions. We begin with the interaction of arginine with itself, which has a very high energy of –35.15 kcal/mol [Table 2, Fig. 1I]. Such stacking interactions are often found in protein structures. Also evaluated was the binding energy for the complex with the two arginine molecules placed face-to-face. Such self-organization is, however, not favored as it leads to charge repulsions (+3.90 kcal/mol). The formation of arginine clusters was studied using quantum calculations and experiments in the gas phase (Julian *et al*. 2001), where it was suggested that cyclic arginine trimers possess exceptional stability. However, in the solvent environment, it was observed that stack interactions also possess good stability. As shown in Table 2, arginine-arginine interactions remain prominent even at room temperature.

Similarly, the parallel stacking interactions for lysine with arginine are also very stable. The binding energy of neutral histidine with arginine is high (–25.89 kcal/mol) due to the presence of a N–H(Arg)···N(His) interaction in the complex (Fig. 1G). The results presented in this section highlight the unique tendency of arginine to form stable clusters in the form of stacks with other arginine molecules. In addition, the results show that arginine also interacts with lysine and histidine through such stacking interactions.

Group 5: Amino acids with sulfur-containing groups

Cysteine plays a crucial role in determining the structure as well as functions of many proteins and has a high reactivity due to the presence of the thiol group. The disulfide bond formed between two cysteines has an important role in protein folding. Sulfur-containing functional groups of cysteine and methionine are normally considered hydrophobic moieties or weak hydrogen bond acceptors in a folded protein structure. Gregoret *et al.* (1991) studied the hydrogen bonds involving sulfur atoms in proteins and found the interaction between arginine and sulfur-containing amino acids to be the N–H···S type. Detailed analyses of these types of S···X interactions (where X = N or O) are

presented by Iwaoka *et al.* (2002). These interactions contribute greatly to the stability of proteins.

The binding energy for cysteine with arginine is –24.88 kcal/mol, while for methionine it is –20.43 kcal/mol. These interactions are similar in nature to those of serine and threonine. Here the hydrogen (–SH) of Cys and the methyl group (–SCH_3) of Met point away from the arginine amino group, while the N–H of arginine forms a hydrogen bond with the sulfur [Fig. 1C]. This hydrogen bond could be due to the $n(S) \rightarrow \sigma^*(N–H)$ orbital interaction in which the N atom tends to approach the lone pair of the S atom. Such a hydrogen bond is usually weaker than the N–H···O type hydrogen bond. The lone N–H pair type interaction could also be hydrophobic in nature since, although sulfur is quite polarizable, it is less electronegative.

Group 6: Amino acids with aliphatic groups

Aliphatic interactions, also termed hydrophobic interactions, play a major role in protein folding and protein aggregation processes. Most of the hydrophobic groups are believed to be exposed to the solvent environment once a protein denatures. Therefore, as arginine is used as a protein stabilizer, the interactions between arginine and the aliphatic amino acids become very important. Although the aliphatic groups do not contain any functional group and thus lack any specific interactions with arginine, favorable interactions of these amino acids with arginine because of induced-dipole interactions were observed [Fig. 2]. The binding energy varies from -15.62 kcal/mol to -18.19 kcal/mol, with the energy of interaction for alanine being the highest and the one for isoleucine being the lowest in Group 6. It is believed that the reason for the observed trend is the larger size of isoleucine, in which the induced dipole gets distributed throughout the molecule. It is notable that the energy for glycine was not computed as it does not have any side chain. Figure 2 shows the induced dipole for the case of alanine. Although the binding energy of the guanidinium group of arginine and aliphatic amino acids is low, it should be noted that the hydrophobic side chain of arginine interacts favorably with aliphatic amino acids through the van der Waals interaction.

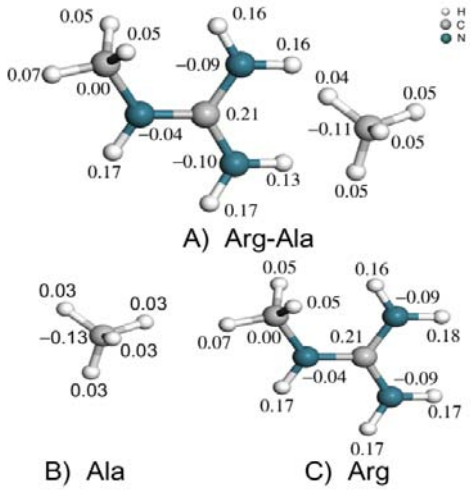

Figure 2. Optimized geometries for A) Arg-Ala complex, B) Arg, and C) Ala. The Hirshfeld charges are also shown. The presence of arginine induces a dipole in alanine.

3.2 Additional Special Interactions of Different Amino Acids with Arginine

In addition to the interactions of arginine with single amino acids, it was observed that arginine can act as a strong link between two amino acids since the side chain of arginine (*ie*. the guanidinium part) has two N atoms that can take part in hydrogen bonding. Selective combinations of arginine with Glu, Asp, Trp, Phe, and Tyr were examined and it was found that arginine can form two strong salt bridges with two acidic amino acids simultaneously. Binding energies for these interactions are -46.33 kcal/mol [Fig. 3A]. In addition, the binding energies for Trp-Arg-Trp, Tyr-Arg-Tyr, and Phe-Arg-Phe complexes were

computed in parallel stacking arrangements [Fig. 3B]. These energies are approximately -10.0 kcal/mol. The T-shape geometry (-30.30 kcal/mol) is more favored than the stacking interactions in the type of bridging interactions above. Hence, the stacking interactions of Arg with aromatic residues are weaker than the salt bridges with the acidic amino acids. These types of interactions can play a key role in protein aggregation.

Next the formation of arginine clusters was probed. As reported in the literature, arginine displays a strong stacking interaction with itself. For example, the DFT studies of Julian *et al.* (2001) suggest that arginine forms cyclic trimer clusters in the gas phase. In these clusters, protonated guanidinium of an arginine interacts with the carboxylate group of another arginine to form stable non covalent complexes in the form of cyclic trimer coordinated to either a cation or an anion. Further, the experimental studies by Das *et al.* (2007) confirmed the formation of arginine clusters and reported that, in a solution environment, hydrophobic interactions (from the aliphatic chain of arginine) play a dominant role in the formation of such clusters. Although hydrophobic interactions arising from the aliphatic side chain of arginine may contribute to the formation of clusters, the current study focuses on the interaction involving the guanidinium group, since, as discussed further below, it is the guanidinium group that is critical to the formation of clusters. Arginine can also form different types of clusters (*eg*. Julian *et al.* 2001), but the focus of the current study is only on the clusters in the form of stacks, as these clusters would be more hydrophobic in nature and might be useful to mask exposed hydrophobic residues on the protein surface.

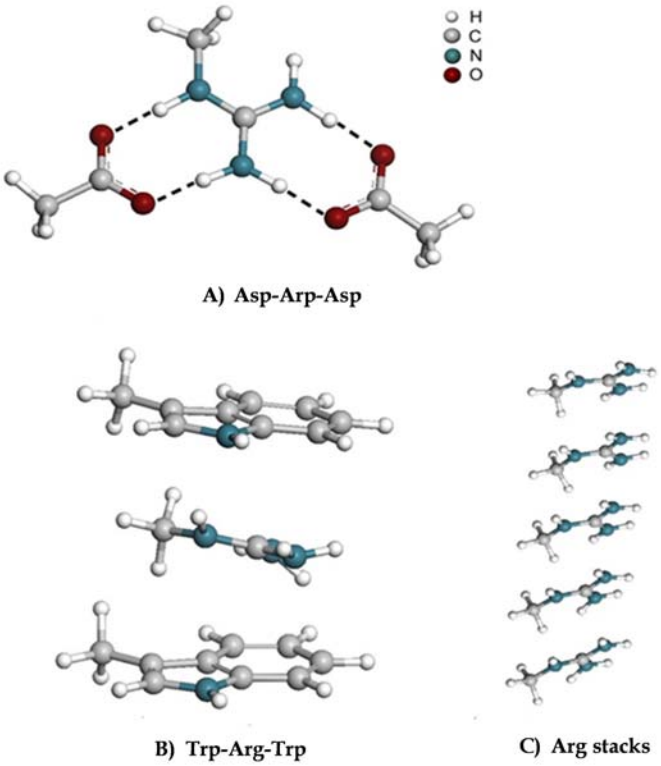

Figure 3. Binding energy (E_b) and relative (to a stack of 2 arginine molecule) LUMO–HOMO (L–H) gap in the arginine stack. All the energies are in kcal/mol.

Table 3. Binding energy (E_b) and relative (to a stack of 2 arginine molecule) LUMO-HOM) (L-H) gap in the arginine stack. All the energies and in kcal/mol.

No. of Arg in stack	E_b	Relative L–H gap
2	–35.15	0.00
3	–49.11	–0.88
4	–63.99	–1.32
5	–75.82	–1.25

The stacking interactions between the guanidinium groups of arginine help in stabilizing the arginine-arginine complex. In order to quantify the optimum arginine pairing in the stacking interactions, DFT calculations for up to five arginine molecules in a stack were carried out; the binding energy increased linearly with the number of arginine molecules and the HOMO-LUMO gap of the complex attains a minimum for a stack of five Arg [Table 3; Fig. 3C]. Das *et al.*[14] suggested that typically 8–10 molecules are present in an arginine cluster. These clusters can be formed by stacking or parallel or both. Calculations in the current study show that clusters containing up to seven arginine molecules are stable based on the binding energy even at 298.15°K [Table 4].

Table 4. Comparison of binding energies between various complexes at 0 K and room temperature (298.15 K).

COMPLEX		ΔE_b (0 K)	ΔG_b (298.15 K)
ASP-ARG-ASP		–46.2	–18.2
TRP-ARG-TRP (stack)		–10.8	+23.5
TRP-ARG-TRP (T-shape)		–30.3	–4.8
HIS-ARG-HIS		–7.7	+22.2
ARG molecules in stack	2	–30.6	–18.3
	3	–41.7	–14.8
	4	–52.8	–12.8
	5	–60.6	–8.3
	6	–71.1	–0.6
	7	–82.4	–0.6

4. Implications to Arginine's Role as a Stabilizer or a Solubilizer

Here the implications and use of the above results in the context of a few typical uses of arginine are discussed.

4.1 Arginine as an Aggregation Suppressant

The results reported above shed light on the mechanism behind arginine's role as an effective suppressant of protein aggregation as calculations clearly suggest that the guanidinium group of arginine interacts with a variety of amino acids. In particular, it forms strong salt bridges with the acidic amino acids and strong stable clusters with itself and can also interact favorably with the aromatic amino acids. Protein structures contain mostly hydrophilic amino acids on the surface, which include Group 1 and Group 4 amino acids, according to the classifications of the current study. The current calculations show that arginine binds strongly with these residues. The aliphatic, methylene groups in the side chains of arginine molecules bound to Group 1 and Group 4 amino acids can easily mask any exposed hydrophobic patches on the proteins, with the polar end-caps on the side chains providing a solvent-friendly shield. In addition, molecular clusters of arginine can also shield the hydrophobic patches on a protein, thereby further reducing the tendency of the proteins to aggregate. The importance of the arginine clusters for stability was also highlighted by Das *et al.* (2007) based on their light scattering and chromatographic analyses of clustering in

arginine solutions and the effect of arginine on heat-induced aggregation of Alzheimer's beta amyloid A$\beta_{1\text{-}42}$ protein. Das *et al.* (2007) suggested that the arginine clusters are stabilized through hydrophobic interactions among the methylene groups. However, while the current researchers' DFT studies also predict clustering of arginine and, in fact, the guanidinium group of arginine contributes to the stability of the resulting clusters. The role of the guanidinium group in the formation of such clusters becomes more evident when one compares the behavior of arginine against that of lysine in solution. Lysine, although quite similar to arginine in structure, lacks the guanidinium group (lysine has an amino group at the end of its side chain), does not form stable clusters, indicating that a similar side chain alone is insufficient to promote the formation of clusters. It is worth noting that the above conclusion concerning the role of the guanidinium group is also consistent with the recent molecular dynamics simulations of Vondrasek *et al.* (2009) on short *di-* or *deca-* arginine peptides. In particular, Vondrasek *et al.*'s (2009) results show significant pairing of the guanidinium groups, which is further confirmed by independent *ab initio* calculations. The results of the current study indicate that the planar geometry of guanidinium promotes its alignment with the guanidinium group of another arginine molecule and causes the formation of arginine clusters in a solution. One might expect, however, that the hydrophobic interactions among the methylene groups might further stabilize the stacks as they are formed, but such interactions alone are insufficient by themselves to sustain clustering. Apart from stacking clusters, as noted in the above calculations, other possibilities of arginine clustering do exist. For example, carboxylate part of an arginine molecule can interact with the guanidinium group of another forming a cluster, but the current researchers believe that only clusters that could create enhanced hydrophobicity (arising from the methylene group of arginine) contribute to protein stability.

In summary, it is proposed that the interactions discussed here contribute to the stability of the proteins in a number of ways. First, the aliphatic side-chains of arginine clusters can mask the exposed hydrophobic patches on the proteins. In addition, the hydrogen bonding interactions between arginine and the hydrophilic residues of a protein, especially the interactions with acidic amino acids, promote the binding of arginine to the proteins (through the guanidinium group), and the aliphatic chains of such bound arginine molecules can also shield any neighboring hydrophobic residues on the protein.

4.2 Thermal Stabilization of Proteins

As noted in Section 1, arginine is also used for the thermal stabilization of proteins. The binding energies for arginine-amino acid interactions reported above in Section 3 shed some light on the potential interactions at play in the thermal stabilization of phytase (Ryu and Park 1998). For instance, the Gibbs free energy for interaction between Arg-Asp is about –17.6 kcal/mol even at 90 °C when the DFT results are adjusted for the effect of temperature (a reduction of about 20% from the magnitude at room temperature), whereas the corresponding magnitudes of Gibbs energies for the interactions of arginine with the other amino acids diminish even more, with some becoming almost zero or even positive. Hence, the ionic interactions between arginine and acidic amino acids might be responsible for inducing stability to these enzymes at high temperatures. The interaction of arginine with various amino acids [Table 2] also identifies the factors that can enhance the stability of proteins in extreme conditions. This is of particular interest because it raises the possibility of engineering enzymes with enhanced high-temperature stability and catalytic efficiency for industrial applications. One of the approaches to achieve this goal would be to use site-directed mutagenesis to replace surface hydrophobic residues with acidic amino residues.

4.3 Arginine as a Solubilizer

Arginine is also widely used to increase the solubility of various proteins, although how arginine enhances the solubility is not known currently. Favorable interactions of amino acid residues with an additive will results in enhancement of the solubility and vice-versa. The current study's DFT calculations adjusted to room temperature suggest that arginine has fairly strong interactions with all amino acids except valine and isoleucine [Table 2]. This is in agreement with the experiments carried out by Arakawa *et al.* (2007) on the solubility of various amino acids in the presence of arginine. Arakawa *et al.* observed that Ile and Val, are indeed less soluble in the presence of arginine, while other amino acids, like Tyr and Trp,

are more soluble. Although an absolute quantitative correlation between the binding energies was not observed in the present study, the solubility enhancement of amino acids as reported by Arakawa *et al.* was calculated; although similar interaction energies of arginine with Try, Trp, and Phe were observed, the solubility of Trp in an arginine solution is higher than that of Phe. The theoretical framework used here can provide an estimate of the solubilizing efficiency of an additive.

The above discussions illustrate that the DFT study presented here not only provides information about important interactions which occur within a protein molecules (*ie.* between the different amino acid side chains and arginine) but also sheds light on the crucial role played by arginine in the suppression of protein aggregation, in increasing the solubility of proteins and thermal stability of important enzymes, and in separation techniques.

5. Conclusions

In summary, the researchers carried out detailed first-principles density functional theory calculations of the structures and the binding energies of the 20 different amino acids with arginine. The theoretical calculations undertaken show that arginine has the strongest interactions with acidic amino acids, followed by interactions with itself. Detailed explanations of the types of interactions and conformation have been discussed above. The binding energies of arginine with all the other amino acids fall in a close range, from -15.62--22.49 kcal/mol (at a ground state) below the binding energies with acidic amino acids. Various interactions such as N–H$\cdots\pi$, a hydrogen bond, van der Waals, induced-dipole, etc. have also been identified as the cause of the stability of the interactions of arginine with the other amino acids. Most of these interactions occur within the protein structure, and the comparative study presented provides a framework for identifying the dominant interactions that help in the stabilization of the protein structure.

As noted in the introductory section of the paper, arginine is a popular choice as an additive for effective suppression of protein aggregation and protein solubilization; however, the mechanistic basis behind the stabilization action of arginine is still not fully understood. The calculations in the current study provide a basis for understanding the role of arginine in this respect. Arginine favorably interacts with the side chains of many amino acids. The results also reveal the strong interaction of arginine with the acidic amino acids through the formation of salt bridges. Acidic amino acids are generally present on a protein surface because of their hydrophilic nature; thus, they are easily accessible to the arginine molecules to interact with. These interactions are believed to be major contributors to increase the solubility of proteins. However, as discussed earlier, which of the numerous other interactions detailed in the paper are also likely to influence stabilization/solubilization depends on the distribution of surface residues and the relative magnitudes of the binding energies in their interactions with arginine. Although a variety of factors influence protein stability and solubility, charge interactions between arginine and acidic amino acids appear to play a major role.

To the best of the authors' knowledge, the present study is the first to examine the interaction between arginine with all other amino acids. In addition, the parallel stacking of arginine with itself as clusters in a solution environment was also analyzed, which would further help in elucidating the functioning mechanism of this unique additive. Finally, it is noteworthy that the approach outlined here can be used for additives other than arginine to analyze systematically how those additives interact with proteins and to design better additives and mixtures of additives.

References

Andzelm J, Govind N, Fitzgerald G, Maiti A (2003), DFT study of methanol conversion to hydrocarbons in a zeolite catalyst. *International Journal of Quantum Chemistry* 91(3): 467-473.

Arakawa T, Ejima D, Tsumoto K, Obeyama N, Tanaka Y, Kita Y, Timasheff SN (2007), Suppression of protein interactions by arginine: a proposed mechanism of the arginine effects. Biophysical Chemistry 127(1-2): 1-8.

Cellamare S, Stefanachi A, Stolfa DA, Basile T, Catto M, Campagna F, Sotelo E, Acquafredda P,

Carotti A (2008), Design, synthesis, and biological evaluation of glycine-based molecular tongs as inhibitors of $A\beta_{1-40}$ aggregation in vitro. Bioorganic and Medicinal Chemistry 16: 4810-4822.

Chakarova-Kack SD, Schroder E, Lundqvist BI, Langreth DC (2006), Application of van der Waals density functional to an extended system: adsorption of benzene and naphthalene on graphite. Physical Review Letter 96(14): 46107-(1-4).

Chen D, Martin ZS, Soto C, Schein CH (2009), Computational selection of inhibitors of Abeta aggregation and neuronal toxicity. Bioorganic and Medicinal Chemistry 17: 5189-5197.

Cleland JL (1991), Mechanisms protein aggregation and refolding. PhD Thesis, Massachusetts Institute of Technology, Cambridge.

Das U, Hariprasad G, Ethayathulla AS, Manral P, Das TK, Pasha S, Mann A, Ganguli M, Verma AK, Bhat R, Chandrayan SK, Ahmed S, Sharma S, Kaur P, Singh TP, Srinivasan A (2007), Inhibition of protein aggregation: supramolecular assemblies of arginine hold the key. *PLoS One* 2(11): e1176.

Delley B (1990), An all-electron numerical method for solving the local density functional for polyatomic molecules. *Journal of Chemical Physics* 92(1): 508-517.

Delley B (2000), From molecules to solids with the DMol3 approach. *Journal of Chemical Physics* 113(18): 7756-7764.

Ejima D, Yumioka R, Arakawa T, Tsumoto K (2005), Arginine as an effective additive in gel permeation chromatography. *Journal of Chromatography* A 1094(1-2): 49-55.

Flocco MM, Mowbray SL (1994), Planar stacking interactions of arginine and aromatic side chains in proteins. *Journal of Molecular Biology* 235: 709-717.

Frokjaer S, Otzen DE (2005), Protein drug stability: a formulation challenge. *Nature Review Drug Discovery* 4: 298-306.

Ghosh R, Sharma S, Chattopadhyay, K (2009), Effect of arginine on protein aggregation studied by fluorescence correlation spectroscopy and other biophysical methods. Biochemistry 48(5): 1135-1143.

Gregoret LM, Rader SD, Fletterick RJ, Cohen FE (1991), Hydrogen bonds involving sulfur atoms in proteins. Proteins Structure Function Genetics 9: 99-107.

Iwaoka M, Takemoto S, Okada M, Tomoda S (2002), Weak nonbonded S•••X (X = O, N, and S) interactions in proteins. Statistical and Theoretical Studies. Bulletin of Chemical Society of Japan 75: 1611-1625.

Ito L, Shiraki K, Matsuura T, Okumura M, Hasegawa K, Baba S, Yamaguchi H, Kumasaka T (2011), High-resolution X-ray analysis reveals binding of arginine to aromatic residues of lysozyme surface: implication of suppression of protein aggregation by arginine. Protein Engineering Design and Selection 24(3): 269-74.

Julian RR, Hodyss R, Beauchamp JL (2001), Salt bridge stabilization of charged zwitterionic arginine aggregates in the gas phase. *Journal of American Chemical Society* 123: 3577-3583.

Klamt A (1995), Conductor-like screening model for real solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *Journal of Physical Chemistry* 99(7): 2224-2235.

Klamt A, Schuurmann G (1993), COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of Chemical Society Perkin Transaction* 2(5): 799-805.

Lee SH, Carpenter JF, Chang BS, Randolph TW, Kim YS (2006), Effects of solutes on solubilization and refolding of proteins from inclusion bodies with high hydrostatic pressure. Protein Science 15: 304-313.

Leippert M, Beck E, Weiland F, Pfaff E (1997), Point mutations within the betaG-betaH loop of foot-and-mouth disease virus O1K affect virus attachment to target cells. *Journal of Virology* 71: 1046-1051.

Lim WK, Rosgen J, Englander SW (2009), Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. Proceedings of Natural Academy of Sciences USA 106(8): 2595-2600.

Martis RL, Singh SK, Gromiha MM, Santhosh C (2008), Role of cation-pi interactions in single chain 'all-alpha' proteins. *Journal of Theoretical Biology* 250: 655-662.

Matsuoka T, Tomita S, Hamada H, Shiraki K (2007), Amidated amino acids are prominent additives for preventing heat-induced aggregation of lysozyme. *Journal of Bioscience and Bioengineering* 103(5): 440-443.

Melo A, Ramos MJ, Floriano WB, Gomes JANF, Leão JFR, Magalhães AL, Maigret B, Nascimento

MC, Reuter N (1999), Theoretical study of arginine-carboxylate interactions. *Journal of Molecular Structure* THEOCHEM 463(1-2): 81-90.

Natsume T, Ishikawa Y, Dedachi K, Tsukamoto T, Kurita N (2006), DFT study of the electronic properties of DNA–DNA and PNA–DNA double strands. *International Journal of Quantum Chemistry* 106(15): 3278-3287.

Ortmann F, Schmidt WG, Bechstedt F (2005), Attracted by long-range electron correlation: adenine on graphite. Physical Review Letter 95(18): 186101-(1-4).

Ottiger P, Pfaffen C, Leist R, Leutwyler S, Bachorz RA, Klopper W (2009), Strong N-H...pi hydrogen bonding in amide-benzene interactions. *Journal of Physical Chemistry* B 113: 2937-2943.

Ryu S, Park TG (1998), Thermal stabilization of aspergillus phytase by L-arginine. Biotechnology Bioprocess Engineering 3(1): 32-34.

Shah D, Shaikh AR, Peng X, Rajagopalan R (2011), Effects of arginine of heat-induced aggregation of concentrated protein solutions. Biotechnology Progress 27(2): 513-520.

Shah D, Li J, Shaikh AR, Rajagopalan R (2012), Arginine–aromatic interactions and their effects on arginine-induced solubilization of aromatic solutes and suppression of protein aggregation. Biotechnology Progress 28(1): 223-231.

Shiraki K, Kudou M, Fujiwara S, Imanaka T, Takagi M (2002), Biophysical effect of amino acids on the prevention of protein aggregation. *The Journal of Biochemistry* 132: 591-595.

Takano K, Higashi R, Okada J, Mukaiyama A, Tadokoro T, Koga Y, Kanaya S (2009), Proline effect on the thermostability and slow unfolding of a hyperthermophilic protein. *The Journal of Biochemistry* 145(1): 79-85.

Timasheff SN (2002), Protein hydration, thermodynamic binding, and preferential hydration. Biochemistry 41(46): 13473-13482.

Vondrasek J, Mason PE, Heyda J, Collins KD, Jungwirth P (2009), The molecular origin of like-charge arginine-arginine pairing in water. *Journal of Physical Chemistry* B 113: 9041-9045.