High-resolution X-ray analysis reveals binding of arginine to aromatic residues of lysozyme surface: implication of suppression of protein aggregation by arginine

Len Ito1,2,5, Kentaro Shiraki3, Takanori Matsuura4, Masaki Okumura2, Kazuya Hasegawa1, Seiki Baba1, Hiroshi Yamaguchi2 and Takashi Kumasaka1

1Japan Synchrotron Radiation Research Institute (SPring-8), 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan. 2School of Science and Technology, Kwansei Gakuin University, 2-1 Gakuen, Sanda, Hyogo 669-1337, Japan. 3Institute of Applied Physics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan and 4Laboratory of Protein Informatics, Research Center for Structural and Functional Proteomics, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

*To whom correspondence should be addressed. E-mail: l-ito@spring8.or.jp

Received May 9, 2010; revised September 15, 2010; accepted October 13, 2010

Edited by Matt Cordes

While biotechnological applications of arginine (Arg) as a solution additive that prevents protein aggregation are increasing, the molecular mechanism of its effects remains unclear. In this study, we investigated the Arg–lysozyme complex by high-resolution crystallographic analysis. Three Arg molecules were observed to be in close proximity to aromatic amino acid residues of the protein surface, and their occupancies gradually increased with increasing Arg concentration. These interactions were mediated by electrostatic, hydrophobic and cation–π interactions with the surface residues. The binding of Arg decreased the accessible surface area of aromatic residues by 40%, but increased that of charged residues by 10%. These changes might prevent intermolecular hydrophobic interactions by shielding hydrophobic regions of the lysozyme surface, resulting in an increase in protein solubility.

Keywords: accessible surface area/arginine/crystal structure/lysozyme/protein aggregation

Introduction

During the past two decades, arginine (Arg) has been used for various biotechnological applications, such as improving the refolding efficiency of recombinant proteins produced in Escherichia coli as inclusion bodies (Buchner and Rudolph, 1991; Brinkmann et al., 1992; Lin and Traugh, 1993; Arora and Khanna, 1996; Hshih et al., 1997; Tsumoto et al., 1998; Bell et al., 2002; Umetstu et al., 2003), solubilizing proteins directly from loose ‘flocculate-type’ inclusion bodies (Tsumoto et al., 2003; Umetstu et al., 2005), eluting antibodies from protein-A affinity resins (Arakawa et al., 2007; Ejima et al., 2005a,b), improving the separation of proteins by column chromatography (size exclusion, ion exchange and affinity) (Ejima et al., 2005a,b) and preventing protein aggregation during refolding and heating (Shiraki et al., 2002, 2004; Hamada and Shiraki, 2007; Matsuoka et al., 2007). In addition, we have determined that the inclusion of certain amino acids and their derivatives promote protein crystallization in highly saturated protein solutions. These additives reduce protein aggregation, resulting in an increased probability of protein crystal (Kobayashi et al., 2007; Ito et al., 2008a,b, 2010a,b). In general, high ionic strength weakens hydrogen bonding and ionic interactions between proteins, and therefore, enhances hydrophobic interactions. In supersaturated protein solutions, high ionic strength may cause amorphous protein aggregation, which can be suppressed by the addition of Arg. The effectiveness of Arg as an additive may be due to its ability to inhibit protein–protein interactions. Despite the widespread use of Arg as an additive, much is to be learned about the mechanism by which Arg decreases protein aggregation.

Crystallographic analysis revealed that the number of hydration water molecules decreases with increasing Arg concentration (Nakakido et al., 2008). In fact, only few Arg molecules bind to lysozyme in aqueous solution, as estimated by preferential interaction and surface tension measurements (Kita et al., 1994; Arakawa et al., 2007; Schneider and Trout, 2009). However, it remains unclear whether Arg molecules directly or indirectly bind to lysozyme. To understand the behavior of Arg on protein surfaces, we determined the crystal structure of lysozyme in the presence of Arg. Due to the weak binding of Arg to lysozyme, the bound Arg residues could only be visualized by high-resolution crystallographic analyses (1.20 Å resolution). The hydrophobic and charged surfaces of lysozyme were compared in the presence and absence of Arg.

Materials and methods

Reagents

All chemicals including Arg hydrochloride were of high-quality analytical grade and were purchased from Wako Pure Chemical Industries (Osaka, Japan). The crystallization sample, a six times crystallized hen egg-white lysozyme (HEWL), was purchased from Seikagaku Co. (Tokyo, Japan).

Turbidity and solubility measurements

A stock protein solution containing 20 mg/ml lysozyme was prepared in 0.1 M sodium acetate (pH 4.5) in the presence and absence of 400 mM Arg or 1000 mM Arg. Stock precipitant solutions with 11 different sodium chloride concentrations (2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6, 3.8, 4.0 and...
4.2 M) were also prepared in 0.1 M sodium acetate (pH 4.5). Precipitant solutions were prepared by mixing 500 μl of the stock precipitant solution in a 1.5-ml microtube. The final solutions contained 1.1–2.1 M sodium chloride, 10 mg/ml lysozyme and 0.1 M sodium acetate (pH 4.5) with and without 200 mM Arg or 500 mM Arg. After 2 days, the turbidity of the mixture was measured at 600 nm using a Jasco spectrophotometer, model V-550 (Japan Spectroscopic Company, Tokyo). After this measurement, all solutions were centrifuged at 15 000 × g for 20 min at 20°C. The concentration of soluble proteins was determined by measuring the absorbance at 280 nm using an appropriate molar absorptivity for lysozyme and the extinction coefficient of 2.63 ml mg⁻¹ cm⁻¹ (Saxena and Wetlauffer, 1970). Reproducibility was evaluated by repeating each experiment more than three times.

**Crystallization**

Crystallization of HEWL in the presence of Arg was performed as follows. A protein solution containing 35 mg ml⁻¹ HEWL in 0, 200 or 500 mM Arg in 50 mM sodium acetate buffer at pH 4.5 was prepared in a 1.5-ml microtube. The protein concentration was determined spectrophotometrically by measuring the absorbance at 280 nm using an appropriate blank solution. The hanging-drop vapor-diffusion method was used for crystallization at 20°C. Hanging droplets were prepared by mixing 1.5 μl of protein solution with an equal amount of a reservoir solution containing 0, 200 or 500 mM Arg, 1.6 M NaCl and 50 mM sodium acetate (pH 4.5). The volume of the reservoir solution was 500 μl for each setting.

**Data collection and determination of crystal structures**

X-ray diffraction data were collected for Arg–lysozyme crystals on beamline BL38B1 at Spring-8 using an ADSC Quantum 210 CCD detector. The diffraction data set was collected under cryogenic conditions from crystals soaked in paratone-N (Hampton Research) and cooled at −173°C in a nitrogen gas stream. High- and low-resolution data were integrated at 4.0–1.2 Å and 50–4.0 Å, respectively, and then merged and scaled using the HKL-2000 program package (Otwinowski et al., 1997). Restrained refinement using the structure of tetragonal lysozyme (PDB-ID 193L; Vaney et al., 1996) in REFMAC (Murshudov et al., 1999) and ARP/wARP v.5.0 (Perrakis et al., 1999) in the CCP4 suite (Collaborative Computational Project Number4, 1994) was used for refinement and to add water molecules to the model. Manual model fitting to the electron density map was performed using the COOT program (Emsley and Cowtan, 2004). Concerning the crystal obtained from the solution containing Arg, cross-validated σA-weighted 2|Fo|−|Fc| and |Fo|−|Fc| electron density maps clearly revealed residual electron densities corresponding to three bound Arg molecules (ARG 1–3). The average temperature factor and occupancy of each Arg was 33.7 and 0.6 Å°, 41.2 and 0.7 Å° and 18.8 and 0.9 Å° for ARG 1, ARG 2 and ARG 3, respectively. Electron density of the main chain atoms of ARG 1–ARG 3 exhibited disorder. The value of the overall average B-factor including protein and water was 14.2 Å°. Root mean square deviations from the ideal geometry were 0.17 Å for bond lengths and 1.8° for bond angles using Sfcheck in the CCP4 suite (Vagunie et al., 1999). Crystal data and relevant statistics are indicated in Supplementary data, Table S1. All figures were prepared using the program PyMOL (DeLano, http://www.pymol.org).

**Intrinsic fluorescence measurements**

Intrinsic fluorescence emission spectra of proteins were obtained at an excitation wavelength of 295 nm with a 10-nm slit width using a RF-1500 fluorescence spectrophotometer (SHIMADZU) at 20°C. Fluorescence spectra were recorded from samples containing 0.3 mg ml⁻¹ protein in 100 mM sodium acetate buffer (pH 4.5). The spectra of samples were corrected by subtracting the corresponding buffers’ spectra.

**Results and discussion**

**Turbidity and solubility measurements**

Effects of 0.2 M Arg on NaCl-induced aggregation were examined using 10 mg/ml lysozyme solution. Figure 1A shows the turbidity at 600 nm as a function of sodium
chloride concentration. In the absence of Arg, little change in turbidity occurred up to 0.9 M NaCl, above which turbidity sharply increased and reached a maximum at 1.2 M. In the presence of 200 mM Arg, a slight increase in turbidity occurred up to 1.1 M NaCl followed by a sharp increase at around 1.5 M NaCl. In the presence of 500 mM Arg, turbidity dramatically increased only at 1.9 M NaCl. These results show that Arg inhibits aggregation of lysozyme induced by NaCl in a concentration-dependent manner.

Comparison of the overall structure of Arg–lysozyme complex with native lysozyme

Cα distances between the crystal structures of the Arg–lysozyme complex and lysozyme alone (PDBID: 193L) are indicated in Fig. 2A. The main chains of these crystal structures had an identical conformation with a RMSD Cα of 0.17 Å. The largest shifts in position of Cα occurred at Val 109 and Ala 110 by 0.77 Å and 0.62 Å, respectively. The other large

---

Table 1. Distances between atoms of ARG1–3 and lysozyme as judged by the distances from 2.1 to 3.5 Å.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Atom</th>
<th>Residue Atom</th>
<th>Distance (Å)</th>
<th>Substrate</th>
<th>Atom</th>
<th>Residue Atom</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG1</td>
<td>N</td>
<td>Wat58</td>
<td>2.93</td>
<td>ARG3</td>
<td>CG</td>
<td>Glu35</td>
<td>3.13</td>
</tr>
<tr>
<td>ARG1</td>
<td>N</td>
<td>Trp62</td>
<td>3.48</td>
<td>ARG3</td>
<td>CG</td>
<td>Wat56</td>
<td>2.48</td>
</tr>
<tr>
<td>ARG1</td>
<td>N</td>
<td>Trp62</td>
<td>3.16</td>
<td>ARG3</td>
<td>CG</td>
<td>Wat56</td>
<td>2.48</td>
</tr>
<tr>
<td>ARG1</td>
<td>CA</td>
<td>Wat60</td>
<td>3.43</td>
<td>ARG3</td>
<td>CD</td>
<td>Glu35</td>
<td>2.62</td>
</tr>
<tr>
<td>ARG1</td>
<td>NE</td>
<td>Asp101</td>
<td>2.83</td>
<td>ARG3</td>
<td>CD</td>
<td>Wat37</td>
<td>2.97</td>
</tr>
<tr>
<td>ARG1</td>
<td>NH1</td>
<td>Trp62</td>
<td>3.14</td>
<td>ARG3</td>
<td>CD</td>
<td>Asp52</td>
<td>3.41</td>
</tr>
<tr>
<td>ARG1</td>
<td>NH1</td>
<td>Trp62</td>
<td>3.34</td>
<td>ARG3</td>
<td>CD</td>
<td>Wat56</td>
<td>2.70</td>
</tr>
<tr>
<td>ARG1</td>
<td>NH2</td>
<td>Asp101</td>
<td>2.96</td>
<td>ARG3</td>
<td>CD</td>
<td>Wat120</td>
<td>3.38</td>
</tr>
<tr>
<td>ARG1</td>
<td>NH2</td>
<td>Wat133</td>
<td>2.75</td>
<td>ARG3</td>
<td>NE</td>
<td>Wat37</td>
<td>2.90</td>
</tr>
<tr>
<td>ARG1</td>
<td>C</td>
<td>Wat60</td>
<td>3.27</td>
<td>ARG3</td>
<td>NE</td>
<td>Wat120</td>
<td>2.89</td>
</tr>
<tr>
<td>ARG1</td>
<td>O</td>
<td>Wat60</td>
<td>2.66</td>
<td>ARG3</td>
<td>NZ</td>
<td>Val37</td>
<td>3.07</td>
</tr>
<tr>
<td>ARG2</td>
<td>N</td>
<td>Wat79</td>
<td>3.25</td>
<td>ARG3</td>
<td>CZ</td>
<td>Val109</td>
<td>3.46</td>
</tr>
<tr>
<td>ARG2</td>
<td>CA</td>
<td>Wat79</td>
<td>2.92</td>
<td>ARG3</td>
<td>NH1</td>
<td>Wat147</td>
<td>3.23</td>
</tr>
<tr>
<td>ARG2</td>
<td>CD</td>
<td>Trp123</td>
<td>3.44</td>
<td>ARG3</td>
<td>NH1</td>
<td>Val109</td>
<td>2.97</td>
</tr>
<tr>
<td>ARG2</td>
<td>NH2</td>
<td>Ala122</td>
<td>2.80</td>
<td>ARG3</td>
<td>NH1</td>
<td>Val109</td>
<td>3.09</td>
</tr>
<tr>
<td>ARG2</td>
<td>NH2</td>
<td>Wat142</td>
<td>3.30</td>
<td>ARG3</td>
<td>NH1</td>
<td>Val109</td>
<td>3.39</td>
</tr>
<tr>
<td>ARG2</td>
<td>C</td>
<td>Wat79</td>
<td>3.29</td>
<td>ARG3</td>
<td>NH1</td>
<td>Wat172</td>
<td>2.60</td>
</tr>
<tr>
<td>ARG2</td>
<td>C</td>
<td>Wat174</td>
<td>3.34</td>
<td>ARG3</td>
<td>NH2</td>
<td>Glu35</td>
<td>3.17</td>
</tr>
<tr>
<td>ARG2</td>
<td>C</td>
<td>Wat44</td>
<td>2.92</td>
<td>ARG3</td>
<td>NH2</td>
<td>Glu35</td>
<td>3.25</td>
</tr>
<tr>
<td>ARG2</td>
<td>O</td>
<td>Wat174</td>
<td>2.11</td>
<td>ARG3</td>
<td>NH2</td>
<td>Glu35</td>
<td>2.56</td>
</tr>
<tr>
<td>ARG2</td>
<td>O</td>
<td>Wat44</td>
<td>2.40</td>
<td>ARG3</td>
<td>NH2</td>
<td>Val109</td>
<td>3.24</td>
</tr>
<tr>
<td>ARG2</td>
<td>OXT</td>
<td>Wat79</td>
<td>3.25</td>
<td>ARG3</td>
<td>NH2</td>
<td>Ala110</td>
<td>3.08</td>
</tr>
<tr>
<td>ARG2</td>
<td>OXT</td>
<td>Wat44</td>
<td>3.36</td>
<td>ARG3</td>
<td>C</td>
<td>Wat120</td>
<td>2.74</td>
</tr>
<tr>
<td>ARG3</td>
<td>CA</td>
<td>Wat120</td>
<td>3.39</td>
<td>ARG3</td>
<td>OXT</td>
<td>Wat120</td>
<td>3.33</td>
</tr>
<tr>
<td>ARG3</td>
<td>CB</td>
<td>Wat36</td>
<td>3.04</td>
<td>ARG3</td>
<td>OXT</td>
<td>Wat132</td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 2 (A) Differences in Cα distances and (B) Difference in main chain relative B-factors between Arg–lysozyme and lysozyme alone (PDBID: 193L). The relative B-factor calculated as follow: mean B-factor of main chain residues divided by mean B-factor of the whole protein. Their superimposition was calculated using the LSQKAB program in the CCP4 suite.
shift in position of Ca atoms was observed in residues His 15–Gly 22 and Gly 67–Arg 73, which are flexible loops at the interface of a and b domains (Hodsdon et al., 1990). The loop region of Thr 47-Asp 48 also exhibits a large shift (Fig. 2A). The difference in average temperature factor of both structures was 24.3 Å² using BVERAGE program in the CCP4 suite (Fig. 2B). The largest differences in main chains relative B-factor between lysozyme alone (PDB-ID: 193L) and Gly 102 by −0.77 and −0.76, respectively. Relative B-factors drop in the complex structures at the three loop regions of His 15–Gly 22, Arg 61–Trp 62 and Gly 67–Arg 73. The regions were stabilized by bound Arg (Table 1).

**Interaction of Arg with lysozyme**

Crystallographic analysis revealed that three Arg molecules were directly bound to lysozyme although with low occupancy and/or high disorder (Fig. 3A). Their occupancies were correlated with Arg concentration because a gradual increase in positive difference electron density was observed with increasing concentrations of Arg (Fig. 3B–J). The primary binding site (ARG 1) occupies the substrate binding site of lysozyme and is formed by the side chains of Trp 62, Trp 63, Asp 101 and Asp 103 (Fig. 3J). The guanidinyl NE and NH2 of ARG 1 forms a hydrogen bond with the side chains OD2 and OD1 of Asp 101, respectively. Binding of this Arg mimics binding of the acetamido group of N-acetyl sugars in GlcNAc complexes with similar chemical interactions (Cheetham et al., 1992). The second binding site (ARG 2) is primarily formed by residues Arg 5, Ala 122 and Trp 123 (Fig. 3G). The guanidinyl NH2 of ARG 2 is involved in hydrogen bonding with the carbonyl O of Ala 122. The third binding site (ARG 3) occupies the active site of lysozyme and is formed by residues Glu 35, Asn 46, Asp

![Fig. 3](https://academic.oup.com/peds/article-abstract/24/3/269/1579339/High-resolution-X-ray-analysis-reveals-binding-of)}
52, Val 109 and Ala 110. The guanidinyl NH2 and CD of Arg form a bifurcated hydrogen bond with the side chain OE2 and OE1 of Glu 35, respectively. There are extensive van der Waal’s contacts between Arg molecules and the side chains or the main chain mentioned above. Interestingly, ARG 1 and ARG 2 also interact with Trp 62 and Arg 5 and Trp 123, respectively, forming unique interactions with the π electrons of their side chains. These so-called cation–π interactions have been originally revealed by analysis of protein crystal structures (Burley and Petsko, 1986). Furthermore, all bound Arg were found near aromatic amino acid residues, such as Trp 62 and Trp 63 (ARG 1), Trp 123 and Phe 34 (ARG 2) and Trp 108 (ARG 3).

Guanidine, an Arg analogue, has similar chemical properties in terms of its structure and high refolding yield. However, Arg is effective in suppressing heat-induced aggregation of lysozyme (Shiraki et al., 2002), whereas guanidine does not suppress this type of aggregation. As described in Arakawa group, binding of Arg is similar to guanidine in the amino acid level, while Arg interacts with the proteins differently from guanidine (Arakawa et al., 2007). Guanidine strongly interacts with hydrophobic amino acids in reduced and denatured lysozyme (Mason et al., 2009), and its sites of interaction with lysozyme complex by X-ray crystallographic analysis have been reported (Mande and Sobhia, 2000). Our finding revealed that Arg binds at sites different from guanidine, and mildly interacts with a nearby hydrophobic amino acid depending on its concentration as described above (Fig. 3) and disrupts intermolecular interactions between thermally unfolded lysozyme molecules by electrostatic interactions. Thus, Arg might act as an aggregation suppressor due to its moderate binding to proteins (Matsuoka et al., 2009). This feature results from a complicated structure through the addition of amino and carboxyl groups to guanidine. These data provide insights into the structural basis of the different mechanisms involved in the suppression of protein aggregation by Arg and guanidine.

**Effects of Arg on intrinsic fluorescence**

The present data showed that fluorescence was partially quenched by addition of Arg. Trp 62 and Trp 63 in lysozyme structure are arranged on one side of the active cleft, Trp 108 is in the active cleft and on the opposite side of Trp 62 and Trp 63, Trp 28 and Trp 111 are intramolecular regions and Trp 123 is located apart from the others. Almost all (80%) of the intrinsic fluorescence of lysozyme is due to Trp 62 and Trp 108 (Imoto et al., 1972). A fluorescence spectrum of lysozyme in 0.1 M Na acetate buffer (pH 4.5) when excited at 295 nm is shown in Fig. 4. The emission intensity by Arg and guanidine increased with an increasing concentration of Arg, suggesting that Arg quenches tryptophan fluorescence, similar to previous reports (Yamashita et al., 1995; Li et al., 2004). This effect may result from the binding of Arg to fluorescent Trp 62 and 108 (Imoto et al., 1972), which is consistent with the structure data (Fig. 3).

**Comparison of accessible surface areas of the Arg–lysozyme complex with native lysozyme**

Table II shows the data for the accessible surface area (ASA) of the Arg–lysozyme complex and average and variance of the native lysozyme structure (PDBID: 1HEL, 193L, 1IEE, 1BWH, 1LJN, 1BVM and 1QTK). The ASA of the Arg–lysozyme complex (6402 Å2) was slightly smaller than that of lysozyme alone (6455 ± 73) due to the binding of ARGs to the small cavities of the protein surface. The contributions of hydrophobic, polar and charged residues to the total ASA were also estimated for the crystal structures (Table II). The ASA of hydrophobic amino acid residues, in particular aromatic amino acid residues, drastically decreased (−40.0%) and the ASA of the charged amino acid residues increased (+13.4%) by in lysozyme/ARG. The ASA of charged amino acid residues includes the ASA of binding ARGs. Basically, the Arg–lysozyme complex exposes less hydrophobic and more electrostatic residues to the solvent.

In supersaturated solutions, high ionic strength causes a decrease in protein solubility (Fig. 1B), which leads to amorphous protein aggregation (Fig. 1A). Our recent observations indicate that Arg reduces protein aggregation during protein crystallization at a high concentration of precipitant (Kobayashi et al., 2007; Ito et al., 2008a, 2010a,b). Our findings support the hypothesis that shielding aromatic residues suppress hydrophobic interactions on the protein surface and...
expose charged residues, consequently solubilizing and stabilizing the protein under high ionic strength solution.

Conclusion
This paper focused on the structural basis of the mechanism by which the addition of Arg decreases protein aggregation and increases protein solubility in a supersaturated protein solution under high salt concentration (or high ionic strength). Although Arg binding to proteins has been implied, this paper reveals clear evidence of such binding using high-resolution X-ray analysis. The structural features of Arg binding are summarized as follows: (i) the bound Arg molecules interact with lysozyme near the Trp residues, typically Trp 62, Trp 108 and Trp 123; (ii) temperature factor for all residues was decreased in the presence of Arg; (iii) the ASA of charged amino acid residues decreases; and (iv) the ASA of charged amino acid residues increases in the presence of Arg. It appears clearly that direct binding of Arg to protein is involved in the aggregation suppression effects of Arg.

Supplementary data
Supplementary data are available at PEDS online.

Funding
This work was supported by the Sasagawa Scientific Research Grant from The Japan Science Society.

References