3.1. INTRODUCTION

The low molecular mass protein i.e., 18 kDa protein was identified in the preputial gland of adult rat and its synthesis was confirmed as testosterone dependent in the previous study (chapter-I). The MALDI-TOF analysis confirmed 18 kDa protein as \( \alpha_{2u} \)-globulin (chapter-II). Further immunohistochemical studies proved that the \( \alpha_{2u} \)-globulin is localized in all acinar cells. The functional significance of this protein is presently speculative. Although it is suspected to serve as a reproductive relevant chemosensory signal, more definitive data needed to support this hypothesis. Therefore, an attempt was made to identify the bound form of volatiles from the purified \( \alpha_{2u} \)-globulin.

In the previous investigation (chapter-I), a testosterone-dependent volatile, farnesol, was found to be present in the preputial gland. It is suspected that this compound shown to have pheromonal activity. Further, this compound may bind with \( \alpha_{2u} \)-globulin for slow release, since rats live in burrows and the pheromonal compound need to persist for longer duration till the conspecifics reach the location. It is also believed that the conspecifics use the pheromones and their binding proteins to trace the pathway of the conspecifics. Since the volatiles and \( \alpha_{2u} \)-globulin were separately identified in the earlier studies, it is unable to justify whether they are bound together or not.

To establish the identification of ligands associated with the protein, it is necessary to purify the protein by chromatographic techniques (gel filtration followed by anion exchange chromatography) and then analyse it by gas chromatography linked
mass spectrometry (GC-MS). This attempt will provide the concrete evidence for the binding of volatiles with the α₂u-globulin.

It is well known that chemical communication in rat is mediated by preputial gland (Gawienowski et al., 1975; Kannan et al., 1998; Zhang et al., 2008; Pohorecky et al., 2008). Eventhough earlier analysis (MALDI-TOF and immunohistochemistry) indicated the presence of α₂u-globulin, confirmatory report is warranted. It can also be done with the help of aminoacid sequencing by Edman degradation protocol. In addition, the amino acid sequences are to be compared with other lipocalins including pheromone carrying proteins of other species to explore the role of α₂u-globulin as pheromone carrier.

For the functional characterization of α₂u-globulin, it needs to be purified for the identification of bound form of volatile compounds. Further, the sequencing of this protein is also important to identify the conserved sequence and to know the sequence similarity with already available lipocalin sequences including pheromone carrying protein. Therefore, the present study focuses the following objectives:

1. to purify the α₂u-globulin by two step chromatography i.e., gel filtration followed by anion exchange chromatography
2. to identify the bound form ligands in the purified α₂u-globulin
3. to obtain the N terminal aminoacid sequence of α₂u-globulin and
4. to check the conserved sequence in the α₂u-globulin and to compare with already available lipocalins in the protein data bank

3.2. MATERIALS AND METHODS
3.2.1. Animals

Adult male laboratory rats (Wistar strain) Rattus norvegicus, were procured from Indian Institute of Science, Bangalore, India, and maintained in the laboratory as mentioned elsewhere in the thesis.
3.2.2. Preparation of the extract

The crude extract from the male preputial gland was prepared by homogenization of the gland in 10 mM Tris/HCl, pH 7.8, using a glass homogenizer, followed by centrifugation at 10,000 rpm for 15 min. The clear supernatant was immediately used in the subsequent purification steps.

3.2.3. Protein purification

In the present study two steps chromatography were carried out for purifying $\alpha_{2u}$-globulin, i.e., i) gel filtration and ii) anion exchange chromatography (Cavaggioni et al., 1990).

i. Gel filtration chromatography

*Principle*

Gel filtration is a form of partition chromatography used for separating molecules of different sizes and is also known as gel exclusion and molecular sieve. The basic principle is that molecules are partitioned between a solvent and a stationary phase of defined porosity. The separation process is carried out using a porous gel matrix (in bead form) packed column and surrounded by solvent. When a mixture of molecules comes in contact with the matrix, the smaller molecules enter the matrix pores, moves slowly and are the last components in the chromatogram. The larger molecules, however, are excluded from the matrix pores and are the first to be eluted. Whereas those intermediate in size can enter the pores but not remain there for long. Thus, all the molecules are eluted in order of their decreasing size.

*Materials and reagents*

1. Column up to 60 cm in length and internal diameter 1.5 cm was used
2. Fraction Collector: Biorad
3. Fraction detector: Spectrophotometer (Perkin Elmer)
Reagents

1. Gel filtration media: Sephadex-G50 (Sigma chemicals)
2. Buffer: 10 mM Tris/HCl, pH 7.8

Procedure

Packing the column

Appropriate amount of gel powder (Sephadex-G50) was suspended in a clean beaker containing 20 volumes of appropriate buffer and was allowed to equilibrate for 1-2 h at 4°C. Sephadex-G50 was gently swirled in the buffer to resuspend the gel. The resuspended gel was then poured carefully and continuously into the glass column, which was kept vertically until the packed portion was just below the top of the column. A conical flask-containing buffer was kept above the level of the column and was connected with the stopcock through a tube, which was attached to the column and it was made sure to be closed. Buffer was added approximately 1/3 volume and the stopcock was closed, maintaining the buffer level above the packed gel (1-2 cm). After setting up the appropriate apparatus, the stopcock was opened after 2 min and the eluted buffer was collected, and approximately 40 ml of buffer was washed through the column as pre-equilibration method. Finally the flow rate was determined as ml/min.

Loading the column

The tubing attached to the column lid and the stopcock was closed. The lid was then removed and it was taped to the buffer bottle. The stopcock was opened and allowed to flow until the buffer head level with the top of the column. It was made sure that the column was not let to run dry. The stopcock was closed and the sample was loaded at the top of the column. The stopcock flow was opened until the sample had moved into the column. The tubing attached to the stopcock was attached to the fraction collector. The fraction collector was then turned on and the fractions were collected. The buffer head was restored carefully by adding the buffer to the top of the column. The column lid was reattached and the stopcock was opened. When the drop counter on fraction collector began to register the drops, the fractions were collected.
**Protein Detection**

For checking the protein availability, each fraction was read at 280 nm.

**Detection of protein purity**

After protein detection, 12 % SDS-PAGE was performed to check the purity of protein as described elsewhere in the thesis.

**ii) Dialysis**

**Principle**

Dialysis is a process, by which small molecules are selectively removed from a sample containing mixture of both small and large molecules. Dialysis is usually performed using a special type of membrane known as semi-permeable membrane. The semi-permeable membrane allows small molecules to pass freely through, holding the large molecules inside. These membranes are essentially made up of cellulose derivatives.

**Materials and reagents**

1. Dialysis membrane (NMWCO- 6-8)
2. 50 mM NaCl in 10 mM Tris/HCl, pH 7.8

**Procedure**

The dialysis membrane (6 cm) was taken and one end of the tubing was tied securely with a thread. Then the fractions containing 18 kDa was filled 2/3 volume using a pipette. The bag was placed in the buffer solution (50 mM NaCl in 10 mM Tris/HCl, pH 7.8) and dialyzed for overnight at 4°C. After the dialysis, the samples were collected carefully and applied to another chromatography step i.e., anion exchange chromatography.
iii) Anion-Exchange chromatography

Principle

Ion-exchange is the separation of proteins on the basis of the net charge which depends on the relative numbers of positive and negative charged groups on the surface of the molecules. The net charge will vary with the pH. The diethylaminoethyl (DEAE) group is used in anion exchange to purify negatively charged proteins.

Materials and reagents

1. Ion-exchanger-DEAE cellulose (Sigma chemicals)
2. Buffer – 10 mM Tris/HCl, pH 7.8
3. Salt solution : 1 M NaCl in 10 mM Tris/HCl, pH 7.8

Procedure

The column was packed with DEAE cellulose (similar to column packing in Gel filtration) and equilibrated with 10 mM Tris/HCl (pH 7.8), approximately 3-4 column volumes. The samples from dialysis were applied on the top of the column. The column was washed with buffer and absorbance was measured at 280 nm under UV-spectrophotometer. The unbound protein was eluted at this stage. Then the column was washed with buffer such that the absorbance reached base line. Bound protein was eluted with a gradient from 0 to 1 M NaCl in 10 mM Tris/HCl (pH 7.8) buffer. Then absorbance was checked at 280 nm and elution continued till absorbance reached base line. After the elution, the purity of each fraction was checked by 12 % SDS-PAGE as per the method described in chapter-I. Then the fractions containing 18 kDa was stored at -20° C for GC-MS analysis.

Extraction of volatiles

For analysis of the small ligands associated with the purified protein, the fraction containing 18 kDa was extracted at room temperature with dichloromethane (1:1).
3.2.4. GC/MS analysis

The GC-MS analysis was performed as per the protocol in chapter-I.

3.2.5. Aminoacid sequence determination

The purified proteins were separated by 12% SDS-PAGE as previously described (chapter-I), electroblotted on poly (vinylidene fluoride) membrane and the N-terminal sequence was analysed with a protein sequencer (Precise 494 HT protein sequencing system, Applied Biosystems).

3.2.6. Comparison studies

The identified sequence of α_{2u}-globulin was compared with lipocalin members using online proteomic tools like multiple sequence alignment programme. The alignment was generated employing CLUSTALW software (Chenna et al., 2003). The sequence similarity of the proteins in the alignment was evaluated consulting a standard 20_20 substitution matrix (Johnson and Overington, 1993).

3.3. RESULTS

The α_{2u}-globulin was purified from preputial gland homogenate using two purification steps such as gel filtration followed by anion exchange chromatography. The result of gel filtration experiments with preputial extract on Sephadex-G50 was shown in Fig. 3.1. This column yielded an improved separation of high and low molecular mass proteins. The fractions from 24 to 32 showed high intensity of 18 kDa proteins. These fractions were pooled and dialyzed against 50 mM NaCl in 10 mM Tris/HCl, pH 7.8. After dialysis, the sample was transferred to anion exchange chromatography. It yielded a purified α_{2u}-globulin after the gradient elution. The purified protein migrated as single band in SDS-PAGE which had a molecular mass of 18 kDa (Fig.3.2).

After purification, the fraction No.11 was extracted with dichloromethane and the extract was analyzed by gas chromatography coupled mass spectrometry for the
identification of endogenous ligands. Four prominent peaks in the gas chromatogram (Fig. 3.3) were clearly attributed to Heptadecane, 2,6,20-trimethyl dodecane, 4 methyl tetradeacne and 3,7,11-trimethyl-2,6,10-dodecatrien-1-ol, after the comparison of their mass spectra of NIST library. The chemical structures of identified compounds were given in Fig.3.4. Totally 31 compounds such as 4-ethyl decane, Heptadecane, Tetradeacne, 2,6,11-trimethyl dodecane, Octadecane, 3,7-dimethyl decane, Nonadecane, 4-methyl tetradeacne, Docosane, 1,1,2,3,3- pentachloro propane, Hexadecane, 2-methyl pentadecane, Eicosane, Octacosane, Pentachlorobutene, 3,5,cyclo-6,8 (14),22-ergostatriene, Squalene, 3-bromocholest-5-ene, 3,7,11,15-tetramethyl-1,6,10,14-hexadecatetraen-3-ol, Geranyl linalool isomer, Oxirane, Farnesol, Tetratriacontane, 2,6,11,15-tetramethyl-hexadeca-2,6,8,10,14-pentene, Cholest-5-en-3-ol (3 beta), Cholestan-3-ol, Cholest-7-en-3-ol, Lanosterol, Tetrakis (2,3-ditert-butyl phenyl)4-4-biphenylene diphosphonate, 2-tert-butyl-4,6-bis (3,5-di-tert-butyl-4-hydroxybenzyl)phenol and 1,2,3-propanetrieyl ester dodecanoic acid were identified in the chapter-I. Among these, four compounds were considered as bound (endogenous) ligands with the α2u-globulin and the remaining compounds were found to be unbound.

The N-terminal amino acid sequence of α2u-globulin contained 10 residues “DDVDKLNGDW”. In order to find out whether GXW in the identified sequence is a lipocalin signature, the identified sequence was compared with other lipocalin members. The lipocalin sequences were obtained from NCBI database and the sequences that have >40% sequence similarities were removed using CD-HIT software. The selected sequences included α2u-globulin, major urinary protein, apolipoprotein D, prostaglandin D synthase, odorant binding protein, major horse allergen, bovide allergen Bos d 2, aphrodisin, salivary lipocalin, retinol binding protein, beta-lactoglobulin, major dog allergens and tear lipocalin (Table 3.1).

In order to find out the information on the secondary structure and location of the identified peptide region in the lipocalin fold, the identified sequence was compared with already available three dimensional structure of α2u-globulin from Rattus
Fig. 3.1. A. Elution profile of preputial gland extract after gel filtration chromatography on Sephadex-G50

B. Fractions were analysed by SDS-PAGE (12%) (M-marker; C-Crude extract; F-Fraction numbers)
Fig. 3.2. A. Elution profile of $\alpha_{2\mu}$-globulin after anion exchange chromatography on DEAE cellulose

B. Fractions were analysed by SDS-PAGE (12%) (L1: Crude extract of preputial gland; L2: Marker; L3: Fraction number 12)
Fig. 3.3. Gas chromatogram of the dichloromethane extract of purified α₂u-globulin from preputial gland of rat
(Peak number refers to 1: Heptadecane, 2: 2,6,11-trimethyl dodecane, 3: 4- methyl tetradecane and 4: Farnesol)
Fig. 3.4. Chemical structures of identified endogenous ligands in the purified $\alpha_{2u}$-globulin
Fig. 3.5. Three dimensional structure of $\alpha_{2u}$-globulin (PDB code:2a2u)

(Alpha helices are marked in blue colour. Beta strands are marked in green colour. Yellow colour shows the loop region)
Table. 3.1. N-terminal sequence of preputial gland $\alpha_{2u}$-globulin compared with those of other lipocalin sequences including pheromone carrying proteins

<table>
<thead>
<tr>
<th>Protein Code</th>
<th>Protein name</th>
<th>Organism</th>
<th>Alignment</th>
<th>Percentage identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha 2u globulin</td>
<td>$Rattus norvegicus$</td>
<td>DDVDKLNGDW</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>109476563</td>
<td>Alpha 2u globulin</td>
<td>$Rattus norvegicus$</td>
<td>LDVDKLNGDW</td>
<td>90</td>
</tr>
<tr>
<td>20178291</td>
<td>Major urinary protein</td>
<td>$Mus musculus$</td>
<td>FNVEKINGEW</td>
<td>50</td>
</tr>
<tr>
<td>6978523</td>
<td>Apolipoprotein D</td>
<td>$Rattus norvegicus$</td>
<td>FDVKKYLGRW</td>
<td>50</td>
</tr>
<tr>
<td>404390</td>
<td>Prostaglandin D synthase</td>
<td>$Homo sapiens$</td>
<td>FQPDKFLGRW</td>
<td>40</td>
</tr>
<tr>
<td>47523248</td>
<td>Odorant binding protein</td>
<td>$Sus scrofa$</td>
<td>QDPFELSGKW</td>
<td>40</td>
</tr>
<tr>
<td>27465627</td>
<td>Odorant binding protein</td>
<td>$Rattus norvegicus$</td>
<td>DDQEDFSGKW</td>
<td>40</td>
</tr>
<tr>
<td>57924235</td>
<td>Apolipoprotein D</td>
<td>$Gallus gallus$</td>
<td>FDINKYLGKW</td>
<td>40</td>
</tr>
<tr>
<td>7766863</td>
<td>Major Horse Allergen</td>
<td>$Equus caballus$</td>
<td>FDISKISGEW</td>
<td>40</td>
</tr>
<tr>
<td>2497701</td>
<td>Allergen Bos d 2</td>
<td>$Bos taurus$</td>
<td>IDPSKIPGEW</td>
<td>40</td>
</tr>
<tr>
<td>4038489</td>
<td>Aphrodisin</td>
<td>$Mesocricetus auratus$</td>
<td>QDFAELQGKW</td>
<td>40</td>
</tr>
<tr>
<td>47523218</td>
<td>Salivary lipocalin</td>
<td>$Sus scrofa$</td>
<td>FDASKIAGEW</td>
<td>40</td>
</tr>
<tr>
<td>40354214</td>
<td>Retinol binding protein</td>
<td>$Homo sapiens$</td>
<td>-MTRDQNGTW</td>
<td>33.33</td>
</tr>
<tr>
<td>119925160</td>
<td>Odorant Binding Protein</td>
<td>$Bos taurus$</td>
<td>QNLELSGSW</td>
<td>30</td>
</tr>
<tr>
<td>126723610</td>
<td>Beta-lactoglobulin</td>
<td>$Equus caballus$</td>
<td>LDLQEVAAGRW</td>
<td>30</td>
</tr>
<tr>
<td>50978938</td>
<td>Major dog allergens</td>
<td>$Canis lupus familiaris$</td>
<td>KDTVAVSGKW</td>
<td>30</td>
</tr>
<tr>
<td>22670</td>
<td>Salivary lipocalin</td>
<td>$Rattus norvegicus$</td>
<td>EENQDVSGTW</td>
<td>20</td>
</tr>
<tr>
<td>4504963</td>
<td>Tear lipocalin</td>
<td>$Homo sapiens$</td>
<td>EEIQDVSGTW</td>
<td>20</td>
</tr>
</tbody>
</table>
Chapter III

novegicus. The three dimensional structure (PDB code: 2a2u) was obtained from Protein Data Bank. As shown in Figure 3.5, the identified sequence falls in the loop and short 3-10 helix region which helps to close off one end of the beta barrel.

3.4. DISCUSSION

The presence of soluble pheromone carrying proteins in secretions such as saliva, sweat, urine and vaginal discharge of several mammals prompted to explore similar substances in the preputial gland of rat. In the present study, the α2u-globulin was successfully purified by two steps of chromatography i.e., gel filtration and anion exchange chromatography. The purified fractions were further processed for GC-MS analysis and Edman degradation which revealed the bound volatiles and aminoacid sequence respectively.

The GC-MS analysis of purified α2u-globulin sample confirmed the presence of four compounds such as heptadecane, 2,6,11-trimethyl dodecane, 4-methyl tetradeclance and farnesol. Among the 31 compounds identified in the intact preputial gland, only four compounds were bound with α2u-globulin. Interestingly, the retention time in which these four compounds appeared in the purified α2u-globulin, was same as that seen in the compounds from the extract of intact preputial gland. Among the four compounds, farnesol has been already suspected to be a pheromone (Kannan et al., 1998).

This approach gains support from many reports on different types of purification methods to identify bound form ligands in the pheromone carrying proteins, for example, MUP secreted from mouse urine and bound with two pheromonal compounds such as brevicomin and thiazole (Robertson et al., 1993). The α2u-globulin has been identified recently in the urine of Indian common house rat (Rattus rattus) which has four bound volatile compounds such as 1-chlorodecane, hexadecane, 2, 6, 11-trimethyl dodecane, and 2-methyl-N-phenyl,2-propenamide. The findings indicate that the α2u-globulin act as a shuttle for pheromone communication in the field.
Aphrodisin, a lipocalin protein abundantly secreted from the vaginal fluid of hamster, facilitates the mating behaviour of the male (Singer et al., 1986; Briand et al., 2000). There are five compounds specially bound onto natural aphrodisin i.e., hexadecanol, 1-octadecanol, z-9-octadecen-1-ol, E-9-octadecen-1-ol and 1-hexadecanol which are also described as part of insect pheromone blends (Briand et al., 2004a,b). Similarly, salivary lipocalin (SAL) secreted from the submaxillary glands of the adult male mature boar contains two endogenous ligands, 5α-androst-16-en-3-one and 5α-androst-16-en-3α-ol as components of the boar sex pheromones (Marchese et al., 1998). The same protein is also expressed in nasal tissue of both sexes in pig, but they are devoid of ligands (Scaloni et al., 2001). In human, the components of axillary odour are associated with apolipoprotein-D and the potent odour is identified as 3-methyl-2-heptenoic acid (Spielman et al., 1995; Zeng et al., 1996). Lipocalin EquCl, abundant protein in horse sweat, contains a ligand, oleamide, which acts as pheromone (D’Innocenzo et al., 2006). All these proteins belong to lipocalin family and possess sequence similarity. Hence, the identified ligands bound to α2α-globulin may act as pheromones.

The testosterone dependent farnesol is a 15-carbon (C15) isoprenoid and belonged to sesquiterpene. It is the dephosphorylated form of farnesyl pyrophosphate, a compound that is common to all branches of the mevalonate pathway and thus precursor of sterols, dolichol, vitamins E and K1, and prenylated proteins. Besides being important biosynthetic intermediate, increasing evidence suggests that farnesol itself plays a direct role in membrane associated biological function (Rowat et al., 2005). Further, it is reported that biosynthesis of pheromones may be strongly associated with the biosynthesis of cholesterol (Rothardt and Beier, 2001). Since the lipocalin is having the capacity to bind with hydrophobic compounds, there is a possibility of farnesol to bind with α2α-globulin. Interestingly, farnesol was already reported as sex attracting compound in rat preputial gland (Kannan et al., 1998). It also mimics the female bee’s sex pheromone in the early spider orchid (Schiestl et al., 2000). Similarly, the compounds EEα farnesene and Eβ farnesene are analogue to farnesol.
reported in mice preputial gland as sex attractant towards female and evoke inter-male aggression (Novotny et al., 1985; Jemiolo et al., 1986).

The identified sequence of $a_2\alpha$-globulin showed homology with several proteins and all of them appear to be members of lipocalin family including pheromone carrying proteins (PCPs). In the alignment, the pheromone carrying proteins i.e., rat urinary $a_2\alpha$-globulin has the highest identity and the two proteins shared 90% of their aminoacids, whereas Major urinary proteins (PCP in mouse) showed 50% similarity. In addition, Major horse allergen (PCP in horse), aphrodisin (PCP in hamster), salivary lipocalin (PCP in boar) were 40% identical to the $a_2\alpha$-globulin of preputial gland. Similarly, the sequence identity of 40% for prostaglandin D synthase (human), odorant binding proteins (rat and boar), apolipoprotein D (hen) and allergen bos d2 (bovine); 33.33% identity for retinol binding protein (human), 30% identity for odorant binding protein (bovine), beta-lactoglobulin (horse) and major dog allergens (dog); and 20% identity for salivary lipocalin (rat) and tear lipocalin (human).

It is important to note that, in the alignment, all the sequences showed the conserved motif i.e., GXW. This motif is usually found within the first 20 residues of all lipocalins and it is almost diagnostic of membership to this family (Sansom et al., 1994; Flower, 1996; Marchese et al., 1998; Thavathiru et al., 1999). Further, the conserved tryptophan (W) has been shown to play a role in maintaining structure, stability and ligand affinity (Gasymov et al., 1998, 1999). Based on the results of the present investigation, it is suggested that the binding site of the protein may be unique in all the lipocalin members.

As $a_2\alpha$-globulin has close sequence similarity with a subset of other mammalian pheromone carrying proteins, the presence of the same in preputial gland might have a pheromone carrying role. Hence, the forthcoming chapter was planned to study the behavioural significance of the farnesol.

66
3.5. SUMMARY

1. The purified protein showed a single band with the apparent molecular mass of 18 kDa.

2. The purified α₂u-globulin still retained the four ligands such as heptadecane, 2,6,11-trimethyl dodecane, 4 methyl tetradecane and farnesol. Among these farnesol alone showed as major peak.

3. N-terminal Edman degradation yielded a single sequence of 10 aminoacids (DDVDKLNGDW).

4. The obtained aminoacid sequence includes the lipocalin signature i.e., -G-X-W- and showed the clear homology with other lipocalins including pheromone carrying proteins.