Removal and Replacement of Endogenous Ligands from Lipid-Bound Proteins and Allergens

Alexander C.Y. Foo[1], **Peter M. Thompson**[1], **Geoffrey A. Mueller**[1]

[1]Nuclear Magnetic Resonance Group, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

Abstract

Many major allergens bind to hydrophobic lipid-like molecules, including Mus m 1, Bet v 1, Der p 2, and Fel d 1. These ligands are strongly retained and have the potential to influence the sensitization process either through directly stimulating the immune system or altering the biophysical properties of the allergenic protein. In order to control for these variables, techniques are required for the removal of endogenously bound ligands and, if necessary, replacement with lipids of known composition. The cockroach allergen Bla g 1 encloses a large hydrophobic cavity which binds a heterogeneous mixture of endogenous lipids when purified using traditional techniques. Here, we describe a method through which these lipids are removed using reverse-phase HPLC followed by thermal annealing to yield Bla g 1 in either its Apo-form or re-loaded with a user-defined mixture of fatty acid or phospholipid cargos. Coupling this protocol with biochemical assays reveal that fatty acid cargoes significantly alter the thermostability and proteolytic resistance of Bla g 1, with downstream implications for the rate of T-cell epitope generation and allergenicity. These results highlight the importance of lipid removal/re-loading protocols such as the one described herein when studying allergens from both recombinant and natural sources. The protocol is generalizable to other allergen families including lipocalins (Mus m 1), PR-10 (Bet v 1), MD-2 (Der p 2) and Uteroglobin (Fel d 1), providing a valuable tool to study the role of lipids in the allergic response.

SUMMARY:

This protocol describes the removal of endogenous lipids from allergens, and their replacement with user-specified ligands through reverse-phase HPLC coupled with thermal annealing. ^{31}P-NMR and circular dichroism allow for the rapid confirmation of ligand removal/loading, and the recovery of native allergen structure.

Keywords

Allergens; Biochemistry; Biophysics; Fatty Acid-Binding Proteins; Fatty Acids; Lipids; Protein Binding; Protein Stability

---

Corresponding Author: Geoffrey A. Mueller, Geoffrey.Mueller@nih.gov, Tel: (984) 287-3589.

INTRODUCTION:

A survey of the allergen database reveals that allergens are found in only 2% of all known protein families, suggesting common functional and biophysical properties contribute to allergenicity.[1] Of these properties, the ability to bind lipid cargoes appears to be strongly over-represented among allergens, suggesting that these cargoes may influence the sensitization process.[1] Indeed, it has been shown that the Brazil Nut allergen Ber e 1 requires co-administration with its endogenous lipid to realize its full sensitizing potential.[2] These lipids could potentially stimulate the immune system directly as illustrated by the mite allergens Der p 2 and Der p 7, both of which share a strong structural homology with LPS-binding proteins.[3, 4] Based on this observation it was proposed that Derp 2 and Der p 7 could bind bacterial lipids and directly stimulate the host immune system through TLR4-mediated signaling, facilitating the sensitization process.[5, 6] It is also possible that endogenously bound lipids could alter the biophysical properties of allergic proteins themselves. For example, the ability of Sin a 2 (mustard) and Ara h 1 (peanuts) to interact with phospholipid vesicles significantly enhanced their resistance to gastric and endosomal degradation,[7] while ligand binding to the major birch pollen allergen Bet v 1 altered both the rate of endosomal processing and the diversity of the resulting peptides.[8] This is particularly relevant to allergenicity given the correlation that has been observed between stability, T-cell epitope generation and allergenicity for proteins such as Bet v 1 and Bla g 1; the latter of which will be the subject of this work.[9, 10]

Bla g 1 represents the prototypical member of the insect Major Allergen (MA) protein family, and possesses a unique structure composed of 12 amphipathic alpha helices which enclose an abnormally large hydrophobic cavity.[9, 11] The available X-ray crystal structure of Bla g 1 shows electron density within this cavity consistent with bound phospholipid or fatty acid ligands; a conjecture confirmed by ^{31}P-NMR and mass spectrometry. These cargoes were heterogeneous in nature and their composition was heavily dependent on the allergen source, with different lipid profiles observed for recombinant Bla g 1 expressed in *E. coli* and *P. pastoris*. Curiously, Bla g 1 purified from its natural allergen source (cockroach frass) contained predominantly fatty acids within its binding site, with a mixture of palmitate, oleate, and stearate being identified as its "natural" ligands.[9, 11] The ability of Bla g 1 to retain lipids and fatty acids following multiple purification steps hinders efforts to study the protein in isolation. Conversely, it has been suggested that the natural palmitate, stearate, and oleate ligands of Bla g 1 (henceforth referred to as nMix) play a key role in both its allergenicity and native biological function.[9] However, these ligands are not present in Bla g 1 obtained from recombinant sources, making it difficult to assess this hypothesis. Similar issues have been observed for other lipid binding allergens such as Bet v 1.[12, 13] To facilitate the systematic study of lipid-allergen interactions we have developed a protocol through which allergens can be quantitatively stripped of their endogenously-bound lipids and reconstituted in either Apo-form, or loaded with specific ligands.

Allergens are most commonly purified from their natural or recombinant sources using affinity chromatography and/or size-exclusion chromatography. Here, we introduce an additional purification step in the form of high-performance liquid chromatography (HPLC) employing a reverse-phase C18 column from which the allergen is eluted into an organic

solvent similar to protocols developed for fatty acid binding proteins.[14] The resulting protein is then subjected to a thermal annealing step in the absence or presence of fatty acids and/or phospholipids. In addition to recovering the native Bla g 1 fold, the elevated temperatures increase the solubility and accessibility of the lipid cargoes, yielding Bla g 1 in either the Apo-form or uniformly loaded with the desired hydrophobic ligand. ^{31}P-NMR spectra of Bla g 1 purified in this manner confirms the complete removal of endogenously bound ligands and uniform replacement with the desired compounds, while circular dichroism confirm the successful recovery of the Bla g 1 fold. The utility of this method is highlighted in a recent work in which cargo binding was found to enhance Bla g 1 thermostability and proteolytic resistance, altering the kinetics of T-cell epitope generation with potential implications for sensitization and allergenicity.[9]

PROTOCOL:

1. **Bla g 1 cloning.**

   1.1. Obtain gene for cockroach allergen Bla g 1.0101 (residues 34–216), representing a single repeat of the MA domain. For the sake of simplicity, Bla g 1 will be used throughout the work to represent this single repeat, rather than the entire Bla g 1.0101 transcript.

   1.2. Subclone the Bla g 1 gene into the desired vector. In this study, the gene containing an N-terminal glutathione S-transferase (GST) tag coupled to a tobacco etch virus (TEV) protease cleavage site was inserted into a pGEX vector for expression as described previously[11].

   1.3. Transform the Bla g 1 pGEX vector into BL21 DE3 *E. coli* cells.

      1.3.1 Prepare a 10 ng/μL stock of the desired vector.

      1.3.2 Combine 1 uL 10 ng/μL DNA stock with 50 μL of BL21 DE3 cells as provided by the manufacturer

      1.3.3 Incubate BL21 DE3-DNA mixture for 30 minutes on ice. Transfer to a 42 °C water bath for 1 minute, then immediately transfer back on ice for an additional 1 minute incubation

      1.3.4 Add 200 μL LB media to the cells and incubate for an additional 1 hour at 37 °C

      1.3.5 Plate the transformed cells on LB-Agar plates containing 100 mg/L ampicillin and grow at 37 °C overnight.

2. **Initial expression and purification.**

   2.1. Inoculate 1 L LB media containing 100 mg/L ampicillin with a single colony of BL21 DE3 cells transformed with the Bla g 1 vector as described in 1.3. Grow at 37 °C overnight.

2.2. On the next day, harvest cells (OD_{600}~1.5) via centrifugation at $6,000 \times g$ for 10 minutes, and resuspend in 2 L 2x YT media containing 100 mg/L ampicillin. Allow cells to grow for an additional 1 hr at 37 °C to an $OD_{600} >0.6$.

2.3. Induce protein expression through the addition of 0.5 mM IPTG. Transfer cells to 18 °C and incubate overnight.

2.4. On the next day, harvest cells as described in 2.2. The resulting cell pellet can be frozen and stored at −20 °C.

2.5. Resuspend pellet obtained from 1 L of culture in 50 mL lysis buffer (50 mM Tris-Hcl pH 8.5, 100 mM NaCl) containing 1 protease inhibitor tablet (or equivalent) and 1 μL benzonase nuclease.

2.6. Lyse cells using a probe sonicator (500 W, 20 kHz) set to 30–50% power for 4 minutes with a 50% duty cycle. Keep the lysate in an ice bath during sonication

2.7. Centrifuge lysate at 45,000g for 20 minutes. Discard insoluble fraction (pellet).

   2.7.1. Remove 28 μL of soluble protein. Combine with 7 μL 5x SDS-PAGE buffer and store for SDS-PAGE analysis. Repeat this step for the GST column flow-through, wash, and elution fractions before and after incubation with TEV.

2.8. Apply soluble proteins (supernatant) to a glutathione resin column (~10 mL total bed volume) equilibrated in PBS pH 7.4.

2.9. Wash out any unbound proteins using 50 mL PBS

2.10. Elute GST-Bla g 1 using 50 mL PBS containing 10 mM reduced glutathione.

2.11. Incubate eluted protein with 0.2 kU TEV protease overnight at 4 °C, or room temperature for 6h to remove GST tag.

3. Endogenous lipid removal via reverse-phase HPLC

3.1. Collect the cleaved Bla g 1 and concentrate it to ~2 mL using a centrifugal filter unit with a <10 kDa molecular weight cut-off.

   3.1.1. Add <12 mL sample to top of concentrator and spin at 5,000 ×g for 10–15 minutes using a swing-bucket rotor.

   NOTE: Sample volume and spin speed will vary based on the specific filter and the type of rotor employed. Consult manufacturer documentation prior to use.

3.2. Load concentrate onto a 250×10 mm HPLC system equipped with a C18 reverse-phase chromatography column equilibrated with 97% buffer A (water, 0.1% trifluoroacetic acid) and 3% buffer B (acetonitrile, 0.1% trifluoroacetic acid).

   NOTE: Smaller columns may be used, but protein may have to be loaded and eluted using multiple cycles to accommodate the reduced binding capacity. When selecting a column ensure that the resin beads have a particle size of < 5

μm and pore size of >200 Å to permit effective separation of protein-sized molecules

CAUTION: Trifluoroacetic acid is highly corrosive, and should be dispensed within a fume hood using appropriate PPE (ie: nitrile gloves, lab coat and goggles).

CAUTION: Acetonitrile is both moderately toxic, volatile, and highly flammable, should be used and dispensed within a fume hood using appropriate PPE (ie: nitrile gloves, lab coat and goggles).

3.3. Elute Bla g 1 using the protocol shown in Table 1 at a flow rate of 1.5–4.0 ml/ min. Monitor elution process using the fluorescence absorbance at 280 nm.

3.3.1. Collect and pool Bla g 1 fractions. Bla g 1 normally elutes at >74% buffer B, or ~34–40 min.

NOTE: Elution time will vary slightly depending on the flow-rate or column size. Collect fractions based on A_{280} for best results.

3.4. Aliquot the sample into glass test tubes, filling no test tube more than halfway (~4 mL). Cover tubes with parafilm and perforate covering with two holes to allow venting.

3.4.1. Prepare a separate 1 mL aliquot (test aliquot). This will be used to determine the expected yield.

3.5. Freeze the samples and test aliquot by placing them in a −80 °C freezer for 1 hr, or immersion in liquid nitrogen. In the case of the later, the tube must be rotated continuously to avoid test tube breakage due to expansion of the liquid phase upon freezing.

3.6. Dry the resulting delipidated protein samples using a lyophilizer. Dried protein may be stored at 4 °C for several months in a sealed container.

4. Reconstitution of Apo- and cargo-loaded Bla g 1

4.1. Determine the anticipated Bla g 1 yield.

4.1.1. Resuspend lyophilized, delipidated (post-HPLC) test aliquot in 5 mL refolding buffer, (50 mM HEPES pH 7.4, 100 mM NaCl, 2% DMSO).

4.1.2. Heat mixture in a water bath (500 mL beaker with 250 mL water and stir bar over a hot plate) to 95 °C. Vortex solutions intermittently and incubate at 95 °C for 0.5–1 hr.

4.1.3. Remove heat and slowly let water bath equilibrate to room temperature (~1 hr). Annealed protein can be stored in this form overnight at 4°C if needed.

4.1.4. Pass annealed Bla g 1-lipid mixture through a 0.22 μM syringe filter to remove particulate matter.

4.1.5. Buffer exchange the filtered protein 3x into PBS pH 7.4 using a centrifugal filter with 10 kDa cutoff as discussed in 3.1 to remove residual free fatty acids and organic solvent.

4.1.6. Assess protein concentration using BCA assay or other preferred method such as OD_{280}. Use this to determine the anticipated yield for the remaining Bla g 1 aliquots.

4.2. Reconstitute Apo- or cargo-loaded Bla g 1.

4.2.1. Resuspend Bla g 1 aliquots in refolding buffer as described in 4.1.1.

4.2.2. To produce Apo-Bla g 1, repeat steps 4.1.2–4.1.6 to obtain desired yield.

4.2.3. To load Bla g 1 with fatty acids, prepare 20 mM stock solutions of the desired fatty acid cargo in methanol or DMSO.

4.2.4. To load Bla g 1 with phospholipids, prepare a 10 mg/mL stock of the desired cargo in chloroform inside a glass test tube.

4.2.4.1. Evaporate the chloroform to produce a lipid film. Add PBS to the test tube to produce a final phospholipid concentration of 20 mM.

CAUTION: Chloroform is harmful if inhaled or swallowed. Use in a chemical fume hood or employ respirator if inadequate ventilation is available. Employ nitrile gloves, lab coat and goggles when handling. Consult MSDS prior to use

4.2.4.2. Rehydrate the lipid film by heating it above the phase transition temperature of the lipid cargo and vortexing until the solution turns cloudy. Note that sonication may be required to fully resuspend and rehydrate some cargoes.

4.2.4.3. If sonication is required, place test tube in bath sonicator (100W, 42 kHz) and sonicate at maximum power until cargo is resuspended. Alternatively, a probe sonicator (described in 2.6) may be used an 10–20% power with a 50% duty cycle.

CAUTION: Sonication employs high frequency sound waves which may damage hearing. Employ noise-suppressing PPE (earplugs or mufflers). If possible, place sonicator inside sound-dampening cabinet or chamber.

4.2.5. Add the desired fatty acid or phospholipid cargo to produce a 20x molar excess of ligands relative to Bla g 1 based on the anticipated

yield determined in 4.1. The total volume of organic solvent added in this step should not exceed 2%. Vortex to mix.

NOTE: 1 L of Bl 21 DE3 cells typically yields ~0.25–0.4 nmol protein, corresponding to ~400 μM ligand per tube.

4.2.6. Anneal the protein as described in 4.1.

5. Confirming phospholipid cargo removal/loading via ^{31}P-NMR

5.1. Concentrate samples of Apo- or cargo-loaded Bla g 1 to >100 μM using a centrifugal filter unit as described in 3.1.

5.2. Rehydrate reference phospholipid in PBS buffer to final concentrations of 2, 1.5, 1, 0.5, and 0.25 mM.

5.3. Dilute samples 1:1 with cholate buffer (100 mM Tris pH 8.0, 100 mM NaCl, 10% w/v cholate) to a total volume of ~600 μL.

NOTE: Cholate is employed in this step to fully extract and solubilize lipids from the Bla g 1 hydrophobic cavity. This ensures that the chemical environment surrounding the phospholipid headgroups is consistent between different samples, allowing for its quantitative assessment using ^{31}P-NMR. The use of cholate can be substituted for chloroform/methanol as described previously.[15]

5.4. Acquire 1D ^{31}P-NMR spectra of the cholate-solubilized Bla g 1 samples and reference phospholipid standards using a broadband probe.

NOTE: The ^{31}P-NMR spectra presented in this work were obtained using a 600 MHz spectrometer. However, previous studies employing similar techniques suggests that acceptable sensitivity can be achieved at fields strengths as low as 150–200 MHz.[15]

5.5. Process the resulting data using appropriate software.[16]

5.6. Obtain peak intensities using preferred NMR viewing software.[17]

5.7. Compare the Bla g 1 ^{31}P-NMR spectra to those obtained for the phospholipid reference samples to confirm removal of endogenously bound ligands and/or binding of desired ligands based on the chemical shifts of the visible peaks (or lack thereof).

5.7.1. Confirm full binding stoichiometry by comparing the peak intensity of the Bla g 1 spectrum to that of the phospholipid reference standards.

6. Confirming Bla g 1 folding:

6.1. Prepare 0.5 μM samples of Bla g 1 in CD buffer (100 mM KH_2PO_4, buffer pH 7.5). Load 1 mL sample into a 10 mm CD cuvette with magnetic stir bar.

6.2. Measure CD spectrum of Bla g 1 to confirm reconstitution of secondary structure. Ensure that HT voltage does not exceed manufacturer recommendations (generally 1 kV).

6.2.1. Measure CD signal from 260–200 nm at 25 °C with a data pitch of 0.2 nm and a scan rate of 20 nm/s with a data integration time of 1 s.

6.3. Increase temperature in the CD cell from 25 °C to 95 °C at a rate of 0.5 °C/min. Activate magnetic stir bar to ensure temperature is uniform across the sample.

6.4. Monitor CD at 222 nm, taking readings every 2 °C.

6.5. Fit resulting data to a 2-state Boltzman curve to determine the melting temperature. Due to the high stability of Bla g 1, the melting temperature (MT_{25}) was defined as the temperature at which the protein has lost 25% of its initial CD at 222 nm.

REPRESENTATIVE RESULTS:

Using affinity chromatography, recombinant GST-Bla g 1 was readily isolated to a high level of purity (Fig. 1A), producing a yield of ~2–4 mg/L of cell culture. Overnight incubation with TEV protease at 4°C is sufficient to remove the GST tag, yielding the final product at ~24 kDa. Note that in this instance there is a significant amount of GST-Bla g 1 in the flow-through and wash fractions, suggesting the Glutathione resin binding capacity was exceeded. The use of more resin or multiple cycles of sample loading and elution could remedy this issue.

Applying the Bla g 1 to a reverse-phase C18 column yields a distinctive elution profile (Fig. 1b), with two large peaks at ~50% buffer B, and a second large peak at ~75% buffer B. SDS-PAGE analysis of the resulting fractions suggest that the former correspond to the cleaved GST tag, while the latter corresponds to Bla g 1. Occasionally a third, smaller peak will occur in the middle corresponding to residual, un-cleaved GST-Bla g 1. The presence of this un-cleaved product can be eliminated by increasing the amount of TEV employed in the cleavage reaction or extending the incubation time. While incomplete cleavage will reduce the yield, the separation obtained on the C18 column is sufficient to ensure that the purity of the final Bla g 1 product remains uncompromised. A consequence of reverse-phase HPLC is that the final protein product is eluted into an organic solvent environment. While this facilitates removal of any hydrophobic ligands, removal of this solvent via lyophilization is required, yielding a fluffy white powder (Fig. 1C).

Annealing of the protein is required to reconstitute the native Bla g 1 fold and can be carried out either in the absence or presence of a lipid cargo. Addition of DMSO to the dried Bla g 1 and phospholipid cargoes prior to the refolding buffer facilitates the solubilization process, though some longer chain lipid cargoes will not fully dissolve even at elevated temperatures. However, this was not observed to impact the loading efficacy among the lipids tested in our studies (Fig. 1C). Similarly, excess lipids will often precipitate out of solution or form large vesicles upon cooling, resulting in a cloudy appearance after annealing (Fig 1C). This was also not observed to effect loading efficiency, and any aggregates are readily removed through the filtration and subsequent buffer exchange steps to yield a clear, transparent solution. Despite the harsh conditions, no thermolysis was observed for Bla g 1.

^{31}P-NMR spectra of Apo-Bla g 1 purified in this manner show no detectable phospholipids either by NMR (Fig. 2A) or thin layer chromotography (data not shown). By contrast, similar spectra obtained for Bla g 1 loaded with a distearoylphosphatidylcholine (DSPC) phospholipid show a strong peak corresponding to the phosphatidylcholine headgroup. For comparison, a representative ^{31}P-NMR spectrum of Bla g 1 purified from recombinant *E. coli* without the use of the lipid removal/annealing protocol described herein (ecBla g 1) show a heterogeneous mixture of endogenous lipids extracted from the recombinant expression system (Fig. 2B). Taking advantage of the quantitative nature of NMR, a standard curve can be produced using reference samples of known DSPC concentrations (Fig. 2C). Comparing the ^{31}P signal intensity obtained from DSPC-Bla g 1 against this standard curve yields a binding stoichiometry of 4.7 ± 0.5 lipids per protein; a value that compares favorably to the predicted full binding stoichiometry obtained from *in silico* studies and structural analysis.[9] Note that this technique will only detect ligands which contain a ^{31}P nucleus such as phospholipids, lysophospholipids, lipopolysaccharides etc. However, this protocol can be easily adapted for ^{13}C-NMR analysis. In this case, methyl-^{13}C labeled fatty acids would be recommended due to its favourable NMR relaxation properties. Restricting isotopic labeling to a single site also facilitates spectral interpretation, as only a single peak is expected, while simultaneously reducing the cost relative to uniform ^{13}C-labeled counterparts. An alternative approach would be to employ mass-spec to identify bound ligands, as demonstrated in a previous study which identified a mixture of fatty acids as the natural cargo of Bla g 1 isolated from cockroach frass (nBla g 1) obtained from commercial sources.[9] However, the limited quantitation capabilities of mass spec precluded an accurate measurement of binding stoichiometry without sufficient standards.

Crystal structures of Bla g 1 reveal a unique fold consisting of 12 amphipathic alpha-helices. Circular dichroism represents a quick and convenient method to assess whether this fold has been successfully reconstituted after the annealing process. CD spectra for Apo- and lipid (nMix)-loaded Bla g 1 show minima ~220 and 210 nm indicative of a predominantly alpha-helical structure (Fig. 3A). This spectrum is extremely similar to that obtained for ecBla g 1 and nBla g 1, providing further evidence that the native structure of Bla g 1 is successfully recovered. This was further confirmed through the use of ^{19}F and ^{1}H-^{15}N solution-NMR, a full discussion of which is available elsewhere.[9] CD-based thermal denaturation assays show a cooperative loss of alpha-helical secondary structure indicative of a folded globular domain (Fig. 3B). Analysis of the resulting melting temperatures (Fig. 3C) show a significant increase upon nMix ligand binding. This elevated thermostability is in line with that calculated for nBla g 1, indicating that we are able to fully reproduce the natural state of Bla g 1. Note that ecBla g 1 also shows a similar, if not greater enhancement in thermostability, illustrating the potential for residual endogenously-bound lipids to interfere with biophysical characterization of allergens purified using traditional FPLC-based approaches. In contrast, the ability to quantitatively remove and re-load hydrophobic cargoes from allergens such as Bla g 1 provides a unique avenue to examine the role of lipids in the allergic response. Here, we describe a method to examine the influence of lipid cargoes on the structure, stability, and endosomal processing of the allergenic proteins themselves, though other avenues of study could be considered.

DISCUSSION:

The protocol described in this work has been successfully applied to systematically study the lipid binding properties of Bla g 1. This revealed a correlation between cargo binding, thermostability, and endosomal processing, the latter of which was correlated with decrease in the generation of a known T-cell epitope with potential implications for immunogenicity. [9, 19] In addition to Bla g 1, other allergens such as Pru p 3 and Bet v 1 have been shown to retain their endogenously-bound cargoes when purified using standard affinity and size-exclusion chromatography methods. [13, 20–22] These unwelcome guests could alter the biophysical and immunological properties of these proteins in a similar manner, highlighting the need for techniques to ensure complete delipidation such as the one presented here.

While the use of reverse-phase HPLC in the purification of allergens has been described previously,[2] coupling it with a thermal annealing protocol provides the rather unusual opportunity to reconstitute allergens with a range of natural and un-natural ligands, allowing users to probe lipid-allergen interactions. This thermal denaturation step was found to be essential for two main purposes. First, thermal denaturation is required to facilitate ligand access to their binding cavities which, due to their hydrophobic nature, are often buried away from the aqueous solvent. [9, 23] Secondly, hydrophobic ligands such as fatty acids and phospholipids often form larger supramolecular structures such as micelles or vesicles when placed in an aqueous environment. The concentration of monomeric, or "free" ligands available for protein binding can be approximated using the critical micelle concentration (CMC). DSPC and other long-chain phospholipids have CMC values in the nM range, indicating that there are virtually no free ligands available for Bla g 1 binding. Even short chain lipids and fatty acids have CMC's in the low μm to mM range, indicating that a large proportion of these ligands remain in the micellar or bilayer phase.[24] However, the high temperatures employed in our denaturation protocol disperses these larger aggregates, facilitating binding. Previous studies have typically employed prolonged incubation periods to facilitate this process. However, the lack of a thermal denaturation/annealing process raises doubts to the efficacy of loading. For example, incubating the mite allergen Der p 5 with the fluorescent fatty acid analogue 11-(Dansylamino)undecanoic acid (DAUDA) yielded a binding stoichiometry of 0.66 despite possessing a large hydrophobic cavity on par with Bla g 1.[25] Likewise, the binding specificity and stoichiometry of plant nsLTPs were found to vary greatly depending on whether the lipids and protein are first solubilized in methanol prior to the addition of aqueous buffer, indicating that ligand and/or binding site accessibility was a limiting factor.[18]

In addition to Bla g 1, we have successfully applied the same strategy to several other MA domain proteins from cockroaches and mosquito (*A. aegypti*), as well as Der p 2 (data not shown). We noted that both the Bla g 1 homologues and Der p 2 eluted at a different time than Bla g 1 from the C18 column (step 3.3). The elution gradients in this step may need to be optimized for other proteins. Alternatively, HPLC columns with a less hydrophobic stationary phase (eg: C8) may be employed, though in the case of Bla g 1 the increased hydrophobicity of the C18 column was necessary to completely remove diacyl phospholipid contaminants from ecBlag 1. Despite the differences in biophysical and biochemical properties we have found this protocol to be extremely robust, and could be easily applied to

other allergenic proteins. While the harsh conditions employed may present a potential limitation, the increased resilience observed for many allergens reduces its impact.[26] Indeed, several food and inhalation allergens such as Der p 2, Ber e 1, Ara a 6 and Lep w 1 have been observed to recover their structure and immunogenicity following thermal denaturation, though optimization of buffer conditions may be required;[27–32] for example reversible denaturation of nsLTP's (Cor a 8) and thaumatins (Mal d 2 and Act d 2) is only observed under acidic (pH <4) conditions.[27, 29, 30] Additionally, it should be noted that the authors did not attempt to optimize either the timing or temperatures employed in the annealing protocol. It is possible that ligand solubilization and protein folding/unfolding may be achieved using a lower maximum temperature as seen with Ber e 1 for which reversible denaturation is achieved at 82 °C.[28] The use of such measures is expected to expand the range of allergens to which this protocol can be applied.

Another important consideration when adapting this protocol to other allergen systems is the concentration of ligands required during the annealing process. In the case of Bla g 1 the expected yield is ~0.25–0.4 μmol of protein per 1 L cell culture. Given the demonstrated binding stoichiometry of 8 fatty acids or 4 diacyl chain lipids per allergen, a 20–40-fold molar excess of cargo (5–10 μmol) was employed. It should be noted that the lipid binding ability of Bla g 1 and its homologues is unique; for example nsLTP's are generally accepted to bind at most two lipid ligands[18] while lipocalins have a stoichiometry of less than 1.[33] As such, complete loading of these types of allergens may be accomplished with a smaller excess of ligands. A final consideration when adapting this protocol to other allergen systems is the presence of disulfide bonds, which can be problematic if not properly formed prior to denaturing. One possible approach would be to carry out the annealing process in the presence of a reducing agent such as 2 mM DTT. The native disulfide bonds could be subsequently re-formed through the addition of reduced and oxidized glutathione as described for the peanut allergen fragment studied by Aalberse *et al.*[34] In this case, recovery of the correct disulfide bonding should be empirically assessed by mass spectrometry.[34]

In this work we describe a technique through which allergens can be delipidated and re-annealed with various phospholipid and fatty acid cargoes. However, there are many other classes of potentially immunogenic or adjuventising ligands present within allergic materials. For example, cat, dog, and mite allergens have been proposed to bind lipopolysaccharides (LPS) and other bacterial lipids from house dust,[35] while the Bet v 1 has been shown to extract complex flavonoids from the pollen matrix.[13] The protocol described in this work can be easily adapted to explore the role of these lipids in a more detailed manner. As a proof of concept we have been able to demonstrate that the hydrophobic cavity of Bla g 1 is capable of binding lipoteichoic acid (LTA) from the cell walls of gram positive bacteria, but excludes LPS from gram negative species, potentially reflecting the greater number of acyl chains in the latter.[9] Taking this one step further, one could utilize the thermal denaturation/annealing protocol to incorporate fluorescent probes and other non-natural fatty acid analogues into allergen proteins. Indeed, we were able to load the hydrophobic cavity of the mosquito homologue of Bla g 1 with DAUDA, opening additional avenues to examine the effects of lipid ligands on allergenic disease.

ACKNOWLEDGMENTS:

REFERENCES:

1. Radauer C, Bublin M, Wagner S, Mari A, Breiteneder H Allergens are distributed into few protein families and possess a restricted number of biochemical functions. Journal of Allergy and Clinical Immunology. 121 (4), 847–852 (2008).

2. Dearman RJ, Alcocer MJC, Kimber I Influence of plant lipids on immune responses in mice to the major Brazil nut allergen Ber e 1. Clinical and Experimental Allergy. 37 (4), 582–591 (2007). [PubMed: 17430356]

3. Reginald K, Chew FT The major allergen Der p 2 is a cholesterol binding protein. Scientific Reports. 9 (1), 1–8 (2019). [PubMed: 30626917]

4. Ling C et al. Expression and refolding of mite allergen pro-Der f1 from inclusion bodies in Escherichia coli. Protein Expression and Purification. 109, 93–98 (2015). [PubMed: 25462803]

5. Trompette A et al. Allergenicity resulting from functional mimicry of a Toll-like receptor complex protein. Nature. 457 (7229), 585–589 (2009). [PubMed: 19060881]

6. Mueller GA et al. The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. Journal of Allergy and Clinical Immunology. 125 (4), 909–917 (2010).

7. Angelina A et al. The lipid interaction capacity of Sin a 2 and Ara h 1, major mustard and peanut allergens of the cupin superfamily, endorses allergenicity. Allergy: European Journal of Allergy and Clinical Immunology. 71 (9), 1284–1294 (2016). [PubMed: 26991432]

8. Soh WT et al. Multiple roles of Bet v 1 ligands in allergen stabilization and modulation of endosomal protease activity. Allergy: European Journal of Allergy and Clinical Immunology. 74 (12), 2382–2393 (2019). [PubMed: 31230350]

9. Foo ACY et al. Hydrophobic ligands influence the structure, stability, and processing of the major cockroach allergen Bla g 1. Scientific Reports. 9 (1), 1–12 (2019). [PubMed: 30626917]

10. Machado Y et al. Fold Stability is a key factor for immunogenicity and allergenicity of the major birch pollen allergen Bet v1.0101. Allergy: European Journal of Allergy and Clinical Immunology. 137 (5), 1525–1534 (2016).

11. Mueller GA et al. The novel structure of the cockroach allergen Bla g 1 has implications for allergenicity and exposure assessment. Journal of Allergy and Clinical Immunology. 132 (6) (2013).

12. Mogensen JE, Wimmer R, Larsen JN, Spangfort MD, Otzen DE The Major Birch Allergen, Bet v 1, Shows Affinity for a Broad Spectrum of Physiological Ligands *. The Journal of biological chemistry. 277 (26), 23684–23692 (2002). [PubMed: 11953433]

13. Seutter von Loetzen C et al. Secret of the major birch pollen allergen Bet v 1: identification of the physiological ligand. Biochemical Journal. 457 (3), 379–390 (2014).

14. Ibáñez-Shimabukuro M et al. Structure and ligand binding of As-p18, an extracellular fatty acid binding protein from the eggs of a parasitic nematode. Bioscience Reports. 39 (7), 1–16 (2019).

15. Beyer K, Klingenberg M ADP/ATP Carrier Protein from Beef Heart Mitochondria Has High Amounts of Tightly Bound Cardiolipin, As Revealed by 31P Nuclear Magnetic Resonance. Biochemistry. 24 (15), 3821–3826 (1985). [PubMed: 2996583]

16. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A Nmrpipe - a Multidimensional Spectral Processing System Based On Unix Pipes. Journal of Biomolecular NMR. 6 (3), 277–293 (1995). [PubMed: 8520220]

17. Johnson BA, Blevins RA NMR View: A computer program for the visualization and analysis of NMR data. Journal of Biomolecular NMR. 4 (5), 603–614 (1994). [PubMed: 22911360]

18. Douliez JP, Michon T, Marion D Steady-state tyrosine fluorescence to study the lipid-binding properties of a wheat non-specific lipid-transfer protein (nsLTP1). Biochimica et Biophysica Acta - Biomembranes. 1467 (1), 65–72 (2000).

19. Dillon MBC et al. Different Bla-g T cell antigens dominate responses in asthma versus rhinitis subjects. Clinical and Experimental Allergy. 45, 1856–1867 (2015). [PubMed: 26414909]

20. Pasquato N et al. Crystal structure of peach Pru p 3, the prototypic member of the family of plant non-specific lipid transfer protein pan-allergens. Journal of Molecular Biology. 356 (3), 684–694 (2006). [PubMed: 16388823]

21. Dubiela P et al. Impact of lipid binding on the tertiary structure and allergenic potential of Jug r 3, the non-specific lipid transfer protein from walnut. Scientific Reports. 9 (2007), 1–11 (2019). [PubMed: 30626917]

22. Abdullah SU et al. Ligand binding to an Allergenic Lipid Transfer Protein Enhances Conformational Flexibility resulting in an Increase in Susceptibility to Gastroduodenal Proteolysis. Scientific Reports. 6 (30279), 1–12 (2016). [PubMed: 28442746]

23. Derewenda U et al. The Crystal Structure of a Major Dust Mite Allergen Der p 2, and its Biological Implications. J. Mol. Biol 318 (1), 189–197 (2002). [PubMed: 12054778]

24. Lipfert J, Columbus L, Chu VB, Lesley SA, Doniach S Size and shape of detergent micelles determined by small-angle X-ray scattering. The journal of physical chemistry. B 111 (43), 12427–38 (2007). [PubMed: 17924686]

25. Pulsawat P et al. The house dust mite allergen Der p 5 binds lipid ligands and stimulates airway epithelial cells through a TLR2-dependent pathway. Clinical and Experimental Allergy. 49 (3), 378–390 (2019). [PubMed: 30230051]

26. Cabrera A et al. Are allergens more abundant and/or more stable than other proteins in pollens and dust? Allergy: European Journal of Allergy and Clinical Immunology. 1267–1269 (2019). [PubMed: 31749155]

27. Offermann LR et al. Structural and Functional Characterization of the Hazelnut Allergen Cor a 8. Journal of Agricultural and Food Chemistry. 63 (41), 9150–9158 (2015). [PubMed: 26417906]

28. Koppelman SJ et al. Reversible denaturation of Brazil nut 2S albumin (Ber e1) and implication of structural destabilization on digestion by pepsin. Journal of Agricultural and Food Chemistry. 53 (1), 123–131 (2005). [PubMed: 15631518]

29. Smole U, Bublin M, Radauer C, Ebner C, Breiteneder H Mal d 2, the thaumatin-like allergen from apple, is highly resistant to gastrointestinal digestion and thermal processing. International Archives of Allergy and Immunology. 147 (4), 289–298 (2008). [PubMed: 18617748]

30. Bublin M et al. Effects of gastrointestinal digestion and heating on the allergenicity of the kiwi allergens Act d 1, actinidin, and Act d 2, a thaumatin-like protein. Molecular Nutrition and Food Research. 52 (10), 1130–1139 (2008). [PubMed: 18655003]

31. Griesmeier U et al. Physicochemical properties and thermal stability of Lep w 1, the major allergen of whiff. Molecular Nutrition and Food Research. 54 (6), 861–869 (2010). [PubMed: 19937851]

32. de Jongh HHJ, de Jong GAH, Apostolovic D, Taylor SL, Baumert JL, Koppelman SJ Effect of heat treatment on the conformational stability of intact and cleaved forms of the peanut allergen Ara h 6 in relation to its IgE-binding potency. Food Chemistry. 326 (April), 127027 (2020). [PubMed: 32438232]

33. Glasgow BJ, Abduragimov AR Ligand binding complexes in lipocalins: Underestimation of the stoichiometry parameter (n). Biochimica et Biophysica Acta - Proteins and Proteomics. 1866 (10), 1001–1007 (2018). [PubMed: 30037780]

34. Aalberse RC et al. Identification of the amino-terminal fragment of Ara h 1 as a major target of the IgE-binding activity in the basic peanut protein fraction. Clinical and Experimental Allergy. 50 (3), 401–405 (2020). [PubMed: 31880850]

35. Bublin M, Eiwegger T, Breiteneder H Do lipids influence the allergic sensitization process? Journal of Allergy and Clinical Immunology. 134 (3), 521–529 (2014).
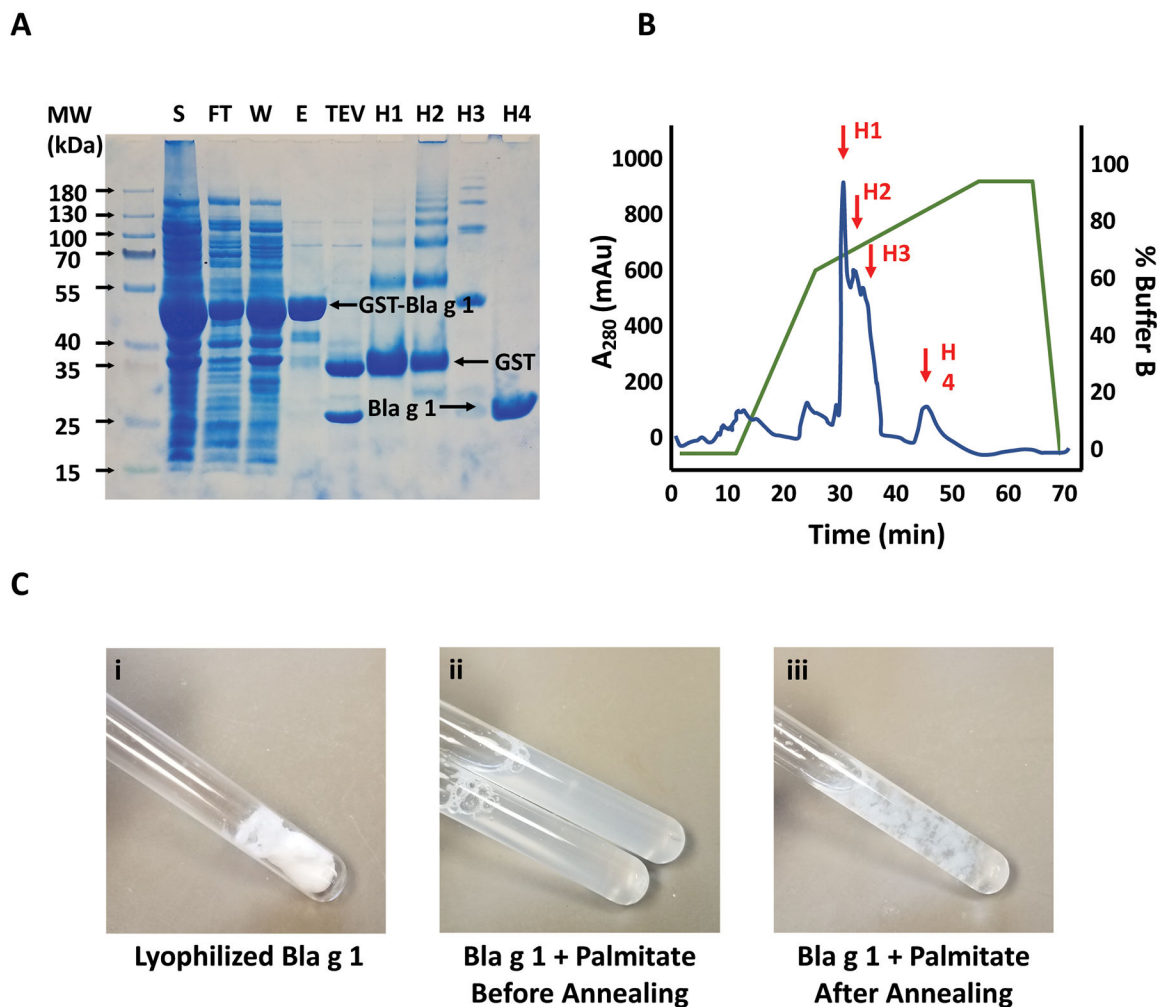
Figure 1: Initial purification of Bla g 1.
A) SDS-PAGE showing the soluble protein fraction following initial lysis (S); flow-through (FT), wash (W), and elution from the glutathione-sepharose column (E); and the final Bla g 1 product following TEV cleavage of the GST tag (TEV). The HPLC elution profile of the resulting Bla g 1 product following TEV cleavage is shown in **(B)**. A_{280} is shown in blue, while the elution gradient (% Buffer B) is shown in green. Fractions corresponding to the cleaved GST tag (H1, H2), residual un-cleaved GST-Bla g 1 (H3), and purified Bla g 1 (H4) are indicated with red arrows at ~50%, ~65%, and ~74% Buffer B respectively. SDS-PAGE analysis of fractions H1- H4 are shown in **(A)** and labeled accordingly. **C)** Representative images showing Bla g 1 at various stages of the annealing process. Note that the precise and extent of precipitate formation as depicted in ii and iii is dependent on the type of lipid cargo employed.
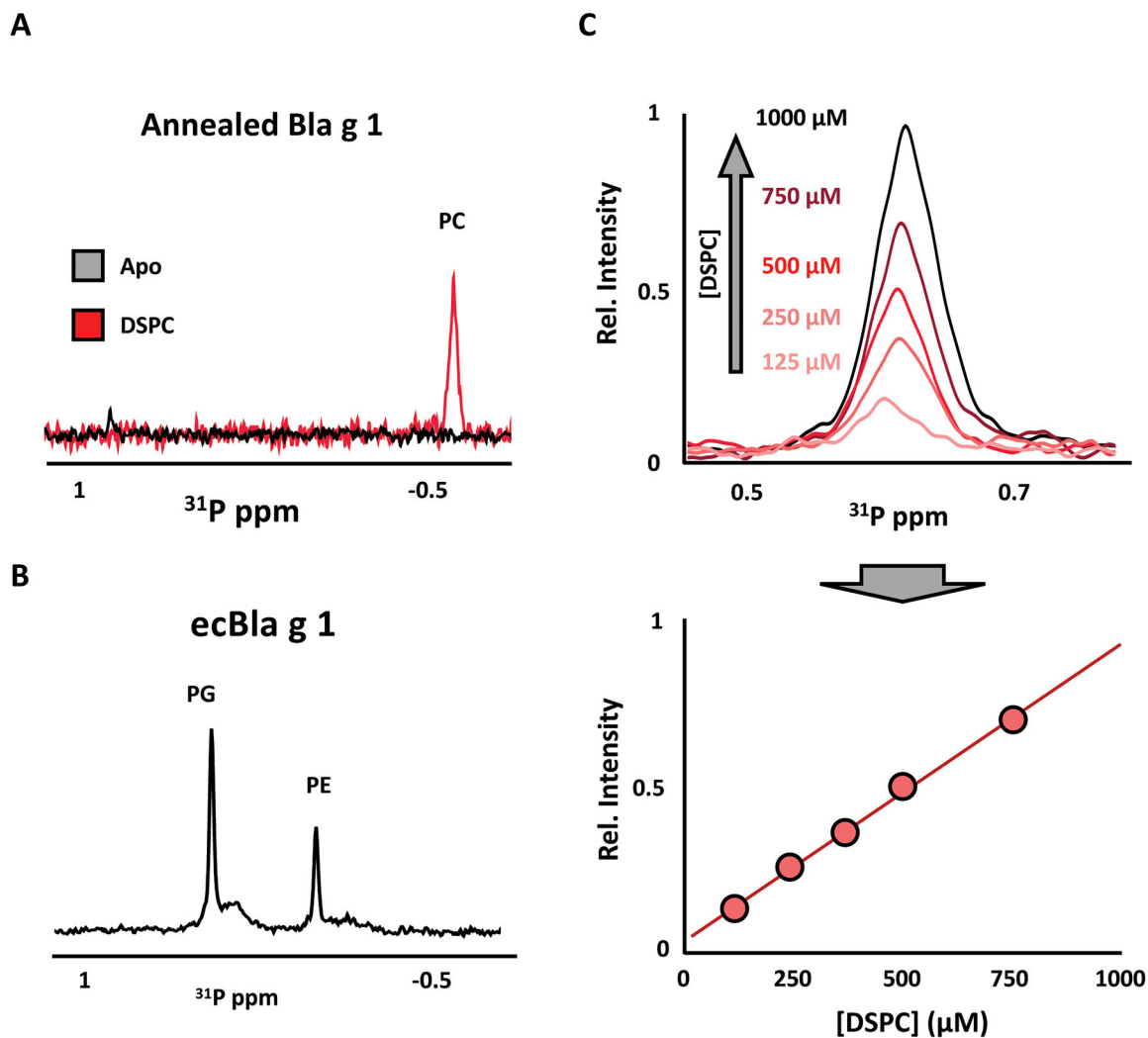
Figure 2: Verifying lipid removal and loading of Bla g 1.

A) ^{31}P-NMR spectra of Apo- (black) or DSPC-loaded Bla g 1 (red) prepared using the annealing protocol described in this work demonstrating the complete removal of lipids in the former, and the homogeneous loading of phosphatidylcholine (PC) lipids achieved in the latter. In contrast, Bla g 1 purified from recombinant *E. coli* without lipid stripping and annealing (ecBla g 1) shows a heterogeneous mixture of endogenous phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) lipids when analyzed using this method **(B)**. A representative standard curve obtained from DSPC reference samples of known concentrations is shown in **(C)**, from which the Bla g 1 binding stoichiometry can be obtained. Figures adapted from Foo *et al.* (2019) and presented under the Creative Commons CC BY License.18
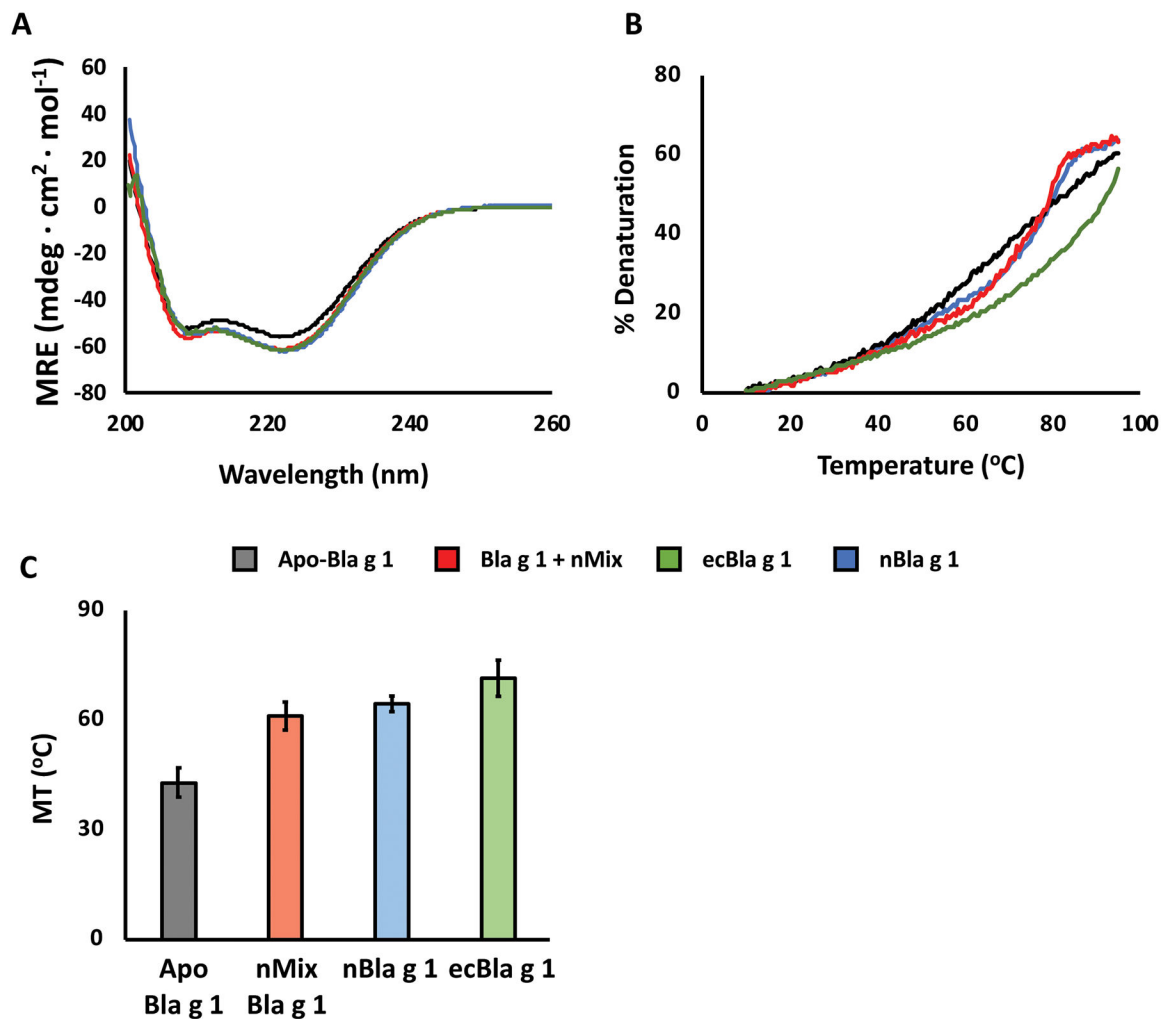
Figure 3: Confirming successful recovery of the Bla g 1 fold.
A) CD spectra of Apo- (black) or nMix-loaded (red) Bla g 1 purified and annealed using the protocol described herein, with minima at ~220 and 210 nm indicative of a predominantly alpha-helical structure consistent with the available X-ray crystal structure. Both Apo- and nMix-loaded Bla g 1 spectra are extremely similar with that obtained for Bla g 1 purified from recombinant *E. coli* (ecBla g 1, green) or from its natural allergenic source (nBla g 1, blue) without the lipid removal and annealing protocol, further supporting the successful recovery of the native structure in the former. **B)** Representative thermal profiles for Apo- (black) and nMix-loaded (red) Bla g 1 showing a sigmoidal curve indicative cooperative unfolding. nBla g 1 (blue) and ecBla g 1 (green) shown as reference. The calculated melting temperatures (MT_{25}) of Bla g 1 are shown in **(C)**. Binding of nMix ligands (red) yields a significant increase in thermostability relative to Apo-Bla g 1 (black). This mirrors the trend observed for nBla g 1 (blue), suggesting that we are able to successfully recover the native state. The even greater stability observed for ecBla g 1 highlights the potential of endogenously-bound lipids to interfere with biophysical characterization of allergens. MT_{25} values presented in C represent the mean value obtained from at least three independent

trials. Error bars represent the corresponding standard deviation values. Figures adapted from Foo *et al.* (2019) and presented under the Creative Commons CC BY License.18

Table 1:

Elution protocol for Bla g 1.

Table illustrating the elution gradient employed in the isolation of Bla g 1 using a C18 HPLC column.

Time (Min)	Buffer A (%)	Buffer B (%)
0	97	3
10	97	3
25	35	65
55	5	95
65	5	95
70	97	3