Published in final edited form as:

Biochim Biophys Acta. 2012 August ; 1824(8): 947–953. doi:10.1016/j.bbapap.2012.04.011.

Mapping the *Anopheles gambiae* Odorant Binding Protein 1 (AgamOBP1) using modeling techniques, site directed mutagenesis, circular dichroism and ligand binding assays

B Rusconi¹, AC Maranhao¹, JP Fuhrer¹, P Krotee¹, SH Choi², F Grun¹, T Thireou³, SD Dimitratos², DF Woods², O Marinotti¹, MF Walter¹, and E Eliopoulos³

¹Developmental Biology Center, University of California Irvine, CA 92697, USA

²Inscent Inc., Irvine, CA 92614, USA

³Department of Agricultural Biotechnology, Agricultural University of Athens, Athens 118-55, Greece

Abstract

The major malaria vector in Sub-Saharan Africa is the *Anopheles gambiae* mosquito. This species is a key target of malaria control measures. Mosquitoes find humans primarily through olfaction, yet the molecular mechanisms associated with host-seeking behavior remain largely unknown. To further understand the functionality of *A. gambiae* odorant binding protein 1 (AgamOBP1), we combined *in silico* protein structure modeling and site-directed mutagenesis to generate 16 AgamOBP1 protein analogues containing single point mutations of interest. Circular dichroism (CD) and ligand-binding assays provided data necessary to probe the effects of the point mutations on ligand binding and the overall structure of AgamOBP1. Far-UV CD spectra of mutated AgamOBP1 variants displayed both substantial decreases to ordered α -helix structure (up to 22%) and increases to disordered α -helix structure (up to 15%) with only minimal changes in random coil (unordered) structure. In mutations Y54A, Y122A and W114Q, aromatic side chain removal from the binding site significantly reduced *N*-phenyl-1-naphthylamine binding. Several non-aromatic mutations (L15T, L19T, L58T, L58Y, M84Q, M84K, H111A, Y122A and L124T) elicited changes to protein conformation with subsequent effects on ligand binding. This study provides empirical evidence for the *in silico* predicted functions of specific amino acids in AgamOBP1 folding and ligand binding characteristics.

Keywords

Odorant Binding Protein; *Anopheles gambiae*; circular dichroism spectroscopy; fluorescence spectroscopy; site directed mutagenesis; ligand binding; 3D modeling

© 2012 Elsevier B.V. All rights reserved.

***Corresponding Authors** –Dr. Elias Eliopoulos, Department of Agricultural Biotechnology, Agricultural University of Athens, Athens 118-55, Greece, Phone: (+30)2105294223, Fax: (+30)2105294322, eliop@aua.gr and Dr. Marika F. Walter, University of California, Developmental Biology Center, Irvine, CA 92697, Phone: (949)824-2185, Fax: (949) 824-2688, mfwalter@uci.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Anopheles gambiae mosquitoes are the main vectors of human malaria in the African continent. Only female mosquitoes feed on blood and are attracted to human hosts for a blood meal primarily by scents [1,2]. The molecular mechanisms of olfaction in insects involve numerous classes of proteins and effectors that interact in the antennal sensilla to translate environmental stimuli into electrophysiological and eventually behavioral responses. Consequently, isolating and understanding the mechanisms of action for particular components of the insect olfactory system provide the basis for rational design of novel mosquito control methods [3].

The *A.gambiae* genome contains about 60 genes encoding putative odorant binding proteins (OBPs) [4,5,6,7]. Among them, *AgamOBP1* is expressed at higher levels in the antenna of females than male mosquitoes and its product has been proposed as a component of the sensory pathway that leads to human host-seeking behaviour [3]. As a result, the three dimensional structure of AgamOBP1 has been determined [8,9] and indole and DEET have been identified as two of its ligands [9,10].

Our understanding of ligand binding by OBPs has been greatly expanded by the availability of the crystal structures of *A.gambiae* OBP1 [8,9] (PDB ID: 2ERB, 3N7H), *Culex quinquefasciatus* OBP1 (CulexOBP1) [11] (PDB ID: 3OGN) and several other insect OBPs. Despite low amino acid sequence conservation among OBPs, their overall structural architecture is similar, revealing elements important for the formation of the odorant binding sites [12]. In this study, we analyzed sixteen single-residue AgamOBP1 mutations, which were predicted by computational simulations to affect protein folding and/or ligand binding. The collected biochemical and structural data provide empirical evidence that validates the predicted functions of several AgamOBP1 amino acids.

2. Materials and Methods

2.1 AgamOBP1 modeling

Three-dimensional (3D) modeling and *in silico* binding studies on AgamOBP1 were based on the crystal structures of AgamOBP1 (PDB code 2ERB and 3N7H) [8,9]. Computational binding studies were performed using AUTODOCK v.4.0 with a Lamarckian genetic algorithm (LGA) search method that simultaneously optimizes ligand conformation and rigid body position [13]. The free-energy scoring function used is based on a linear regression analysis, the AMBER force field, and on a large set of diverse protein-ligand complexes with known binding constants.

Topohydrophobic (TH) residues are highly conserved hydrophobic residues found among very divergent members of a protein family. They are hydrophobic amino acids belonging to the folding nucleus which are rapidly protected during folding and play an important role in the first stages of protein folding [14–16]. In order to determine TH positions, structure-based sequence alignment of typically five or six members of a protein family sharing low sequence identity is carried out. A position in this alignment is considered as TH if at least 75% of the residues found in this position are strong hydrophobic amino acids (VILFMWY) and the remaining residues belong to the ACTQERKH group [16]. TH residues of AgamOBP1 were determined following the structure-based sequence alignment of crystallographically determined structures of the insect pheromone/odorant-binding proteins SCOP family [17,18] (PDB code 2ERB, 3CZ1, 1OOH, 1OW4 and 1DQE) using the Dali server [19].

AgamOBP1 protein structure was also analyzed as a succession of Tightened-Ended-Fragments (TEFs) [20,21]. TEFs are amino acid sequences (fragments) having their ends located in the protein core and close to each other in the protein tertiary structure (C-alpha carbon distance smaller than 10Å). Most of the TEF limits are buried in the hydrophobic protein core and represent good candidates for involvement in the folding nucleus. TEF length varies typically from 10 to 50 residues (mean length of 25 residues). TEF determination was carried out at the RPBS web server (<http://bioserv.rpbs.univ-paris-diderot.fr/TEF>) [22]. TEF calculation for AgamOBP1 was based on the PDB structure 2ERB, setting the maximum distance between the alpha carbons of the two amino acid residues on the limits of a TEF equal to 10Å. TEF minimum and maximum lengths were set at 10 and 25 residues respectively.

2.2 Plasmid construction and expression and purification of recombinant AgamOBP1

The AgamOBP1 (AF437884) cDNA, without the sequence encoding the secretory signal peptide, was cloned in pRSET-B [10]. The recombinant His-tagged fusion protein (rAgamOBP1) was expressed in *E. coli* BL21 Star (DE3)pLysS cells in a soluble form and was purified on a nickel-NTA column following the manufacturers' directions (Thermo Fisher Scientific), eluted with 5mM EDTA and subjected to extensive dialysis against 50 mM Tris-HCl pH 7.4.

2.3 Site directed mutagenesis

The QuikChange site-directed mutagenesis kit (Stratagene) was used to generate point mutations in the pRSET-B cloned AgamOBP1 cDNA (Table 1). Oligonucleotide primers, each complementary to opposite strands of the vector, were designed so each mutation changed a single amino acid in the sequence of the encoded rAgamOBP1. All plasmids containing mutations were verified by DNA sequencing. Mutated protein expression and purification followed the same protocol described above for rAgamOBP1, except when the mutations resulted in accumulation of the recombinant proteins in inclusion bodies. In these instances, inclusion bodies were collected by centrifugation and the insoluble proteins were dissolved in 9M urea - 50 mM Tris pH 8.0 buffer. The protein was solubilized in 8M Urea and 5mM TCEP-HCl as reducing agent, was bound to the Ni column, and refolded on-column using a urea gradient [23] (Rehm et. al., 2001). After elution, the refolded protein was extensively dialyzed against 50 mM Tris-HCl pH 7.0.

2.4 Circular Dichroism (CD)

The CD spectra of rAgamOBP1 and rAgamOBP1 mutants were recorded at 20°C using a Jasco Model J-810 spectropolarimeter (Japan Spectroscopic Co., Tokyo, Japan). Samples were scanned at protein concentrations of 0.2 – 0.5 mg mL⁻¹ in 10 mM sodium phosphate, pH 7.5. CONTIN [24] was used as deconvolution algorithm and reference sets contained CD spectra of globular proteins of known X-ray crystal structure, representative of α -helix, β -strand, β -turn and unordered elements[25].

2.5 In vitro binding assays

Ligand binding assays [26] were carried out for rAgamOBP1 or its mutated forms using 1-NPN (N-Phenyl-1-Naphthylamine, CAS 90-30-2), a fluorescent dye that modifies its emission spectrum upon binding the insect OBPs. A ligand, 3 methyl indole, capable of binding AgamOBP1 via the protein's binding pocket was used to competitively displace 1-NPN and in doing so quenching fluorescence. The assays were conducted with 8 μ M 1-NPN, 16 μ M 3 methyl indole and 4 μ M of rAgamOBP1. Fluorescence was detected using a Molecular Devices Gemini XPS spectrofluorometer (Sunnyvale, CA).

For each sample tested, the area under specific regions of the fluorometry response curve was used to quantify the 1-NPN binding to AgamOBP1. To calculate this area, we took into consideration data collected from 400 nm to 440 nm and exceeding 200 RU (response units).

3. Results

3.1 Three-dimensional modeling of AgamOBP1 and the design of AgamOBP1 mutants

Computer assisted molecular modeling has been widely used in drug design and discovery of bioactive compounds, including insect repellents [27]. Amino acid residues putatively involved in AgamOBP1 structure maintenance and ligands binding were identified by similar *in silico* studies (Figure 1). Sixteen mutated forms of AgamOBP1 (Table 1), each representing a single amino acid substitution designed to affect protein folding and/or ligand binding, were expressed and the predictions evaluated.

3.1.1. Mutations associated with positions affecting AgamOBP1 folding—

Topohydrophobic (TH) residues differ from non-conserved hydrophobic amino acids because they are generally buried in the inner part of the protein hydrophobic core, have low side-chain dispersion (the orientation of the side-chains in the same position is conserved among proteins of the same family), form a network of directly interacting residues and show good correlation with the residues essential for the folding process [14–16]. Consequently, mutation of TH residues is likely to alter protein folding [28].

TH positions were identified by structure-based sequence alignment of five pheromone/odorant-binding proteins. The identified TH amino acids were L15, L19, I37, H46, Y54, M55, L58, V64, V65, V71, H77, L80, M84, F108, L110, and W114 (Numbering from OBP1 structure PDB code 2ERB). TEF calculations for OBP1 were based on the PDB structure 2ERB, and the predicted TEF limits were V32, K50, V64, K75, H77, A88, M91, H111, W114, and Y122. Approximately 50% of the TH residues are found near (± 3 residues) the TEF limits (Figure 1). Mutations that potentially affect AgamOBP1 folding were designed and performed on or near the predicted TH positions and / or TEF limits (± 2 residues) (Figures 1A, 1B, Figure 2, Table 1). Several of these amino acids are also involved in the ligand binding site lining as described below (59, 84, 114, 122, 123, 124) (Figure 1C).

CD analysis was used to evaluate changes in protein structure caused by mutations. The far-UV CD spectrum for the wild type AgamOBP1 exhibits a profile with a maximum at 192 nm and a negative intensity over a relatively wide region with minima at 208 nm and 225 nm, characteristic of the contribution of α -helical domains within the protein (Figure 3) [29]. In contrast, the far-UV CD spectra of all but one (F123Y) of the AgamOBP1 mutant proteins exhibited a significant disruption of the overall secondary structure of the proteins (Figure 3). In all mutant proteins, the intensity of the 190 – 198 nm spectral maxima decreased and the intensities of the 225 nm minima increased suggesting a possible unwinding of an α -helix structure and a corresponding increase in random coil content.

All mutationally-induced perturbations to the secondary structure of AgamOBP1 protein can be divided into four distinct spectral types based on the CD data obtained (Table 1, Figure 3). Type I is represented by the mutant in which residue 123, a non-polar phenylalanine, is replaced by a polar tyrosine (F123Y), resulting in a minimal change from the characteristic α -helix -rich AgamOBP1 CD pattern; a slight loss of intensity in the region 190 – 198 nm, no change of the minimum at 208 nm, and a slight loss of the minimum at 225 – 226 nm. The Type II pattern demonstrates the effects of the substitution of a non-polar methionine for the polar glutamine at residue 84 (M84Q). This mutation results in a dramatic loss of UV intensity at the 190 – 198 nm maximum and a considerable loss of the 225 – 226 nm minimum. However, the 208 nm minimum is retained as in Type I. Ten mutants produced

CD spectra classified as Type III (L15T, L19T, L58Y, Y54F, M84K, H111A, Y122A, Y122F, F123T, L124T). These mutations resulted in spectra that exhibited a major loss of intensity in the 190 – 198 nm region, a shift of the 208nm minimum to 206nm and a significant decrease in the minimum at 225 – 226 nm. Type IV spectra produced by mutations W114Q, L58T, Y54A, and F59Y, in or near TH positions, demonstrate a significant loss of intensity in the 190 – 198 nm region and in the 225 – 226 nm region as well as the largest shift in the 208 nm minimum to 204 nm. All of these spectral changes suggest either loss of α -helix structure and a corresponding increase in random coil content or a reorientation of the respective regular α -helical elements.

In order to estimate the secondary structure of OBP1 and the 16 OBP1 mutant proteins, deconvolution of the CD spectra in the far-UV region was performed and the secondary structure calculations are summarized in Table 2. We considered that if the NRMSD is less than 0.2, the calculated structure is characterized as having a secondary structure similar to the actual structure. If the NRMSD is equal to or greater than 0.2, the calculated structure does not resemble the actual structure [30]. The NRMSD values for all but two mutants were less than 0.2 and six were less than 0.1, indicating a good fit of the experimental and calculated data and thereby, a relatively accurate analysis of secondary structure.

These data suggest that the native AgamOBP1 protein contains more than 65% regular α -helical structure and less than 2% regular β -strand structure. In mutated AgamOBP1 proteins, ordered α -helical structure (Helix 1) decreased by as much as 22% as a result of the amino acid substitutions. In all but two mutants, disordered α -helical structure (Helix 2) correspondingly increased. In nearly all cases a slight decrease in β -strand secondary structure occurred and in several mutants increases in β -turn structure occurred. Eight mutants showed increases in unordered secondary structure.

3.1.2. Mutations associated with the binding site—The AgamOBP1 binding cavity has an elongated cylindrical shape with the wider part being towards the rim of the entrance and running through the monomer. It is formed by amino acid residues belonging to helices α 4 (Leu73, Leu76, His77, Leu80), α 5 (Ala88, Met89, Met91, Gly92), α 6 (Trp114), lined mainly with hydrophobic residues (Leu15, Leu19, Leu58, Phe59, Leu76, Leu80, Met84, Leu124), as well as other residues possessing polar properties (His111, Trp114, Tyr122); (numbering as in pdb code 2ERB) (Figure 1B). The binding pocket is wide enough to accommodate 1-NPN, bearing two aromatic subunits (Figure 1C) as well as other molecules without long side chains, such as oleic acid or polyethylene glycol (PEG). Proteins mutated at positions, 59, 84, 111, 114, 122, 123, 124, were designed to alter physicochemical properties within the AgamOBP1 binding site. In vitro binding assays performed with the mutated proteins demonstrated changes in the binding affinity of the protein thus validating the predictions (Figure 4 and Supplemental Figure 1). Mutations H111A, M84Q, W114Q, Y122F and F123Y did not affect protein solubility and resulted in reduced 1-NPN binding, suggesting that the native residues are involved in ligand binding. Mutations, F59Y, M84K, Y122A, F123T and L124T, altered protein conformation and reduced ligand binding. Mutations that did not significantly alter amino acids properties in the predicted binding site resulted in a minimal effect on 1-NPN binding. (i.e. at positions 84, 122, 123). Less conservative changes affecting the predicted 1-NPN binding site at positions 59, 84, 111, 114, 122, 123, and 124 significantly weakened 1-NPN binding (Figure. 4, Supplemental Figure 1).

4. Discussion

We modeled the binding of 1-NPN, a generic ligand for insect OBPs, to AgamOBP1 and predicted amino acid involved in the binding site formation as well as protein folding and

assembly. The predictions were evaluated and validated by CD and ligand binding assays, as follows.

The far-UV CD spectra of wild type AgamOBP1 obtained in this study demonstrate a characteristic α -helix-rich protein profile. CD spectra of all but one of the OBP1 mutant proteins (F123Y) exhibited decreases of ordered α -helix structure of 3 to 22 percent and increases of disordered α -helix structure of 4 to 15 percent but surprisingly, only minimal changes in random coil (unordered) structure. Significant increases of β -turn structure of 3 to 15 percent were calculated for twelve of the mutant proteins. The CD spectra of most of the mutant proteins also exhibit shifts towards the far UV, mainly the 208 nm to 205 nm shift as well as differences in the amplitude of the 190 – 198 nm and the 225 – 228 nm regions as a result of loss of regular α -helical structure and increase of disordered α -helix content [31]. Substitutions of aromatic residues (Tyr or Trp) for non-aromatic ones may result in significant difference in the spectral shape since these amino acids may contribute significantly to the far-UV spectrum [32]. The introduction of hydrogen bond potential by substitution of polar residues may also affect the far UV spectrum by altering (potentially either stabilizing or destabilizing) the ionized side chains that interact with other secondary structural elements in the protein [33].

Our results demonstrate that mutations Y122F and F123Y, which cause no significant change to the residue aromatic properties, had an insignificant effect in 1-NPN binding or the solubility of the protein. Loss of a hydroxyl group in the TEF limit position 122 significantly shifted the CD spectrum towards disordered α -helix indicating increase of disorder. Introduction of a hydroxyl at position 123 located in the binding site moderately lowered 1-NPN binding without altering the CD spectrum, thus showing that a minor alteration of binding properties occur with no significant change in protein conformation.

Aromatic side chain removal (Y54A, Y122A, W114Q), resulted in significant changes in CD spectra and major reductions in 1-NPN binding indicating the contribution of the aromatic side chains both in the far UV spectrum and ligand binding properties. Mutations L15T, L19T, L58T, L58Y, M84Q, M84K, H111A, Y122A and L124T all produced significant changes in the CD spectrum, indicating changes in protein conformation with subsequent effects on binding, depending on the position of each amino acid, in or out of the binding site.

Site-directed mutagenesis, together with CD and ligand binding studies provided empirical evidence for the *in silico* predicted functions of specific amino acids in AgamOBP1. This data provide a valuable insight into the 3D structure and function of AgamOBP1 and combined with our work on mutagenesis and ligand binding on other OBPs of the same species, assist our effort in understanding mosquito olfaction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge the late Dr. Harald Biessmann (Developmental Biology Center, University of California Irvine), for kindly providing the AgamOBP1 gene. This work was supported by funding from the European Commission for the FP7- HEALTH-2007-2.3.2.9 project “ENAROMaTIC” and the US National Institutes of Health for grants R56AI05148-6A1 and RAI051485..

References

1. Foster WA. Mosquito sugar feeding and reproductive energetics. *Annu. Rev. Entomol.* 1995; 40:443–474. [PubMed: 7810991]
2. Takken W, Knols BG. Odor-mediated behavior of Afrotropical malaria mosquitoes. *Annu Rev Entomol.* 1999; 44:131–157. [PubMed: 9990718]
3. Justice RW, Biessmann H, Walter MF, Dimitratos SD, Woods DF. Genomics spawns novel approaches to mosquito control. *Bioessays.* 2003; 25:1011–1020. [PubMed: 14505368]
4. Vogt RG. Odorant binding protein homologues of the malaria mosquito *Anopheles gambiae*; possible orthologues of the OS-E and OS-F OBPs OF *Drosophila melanogaster*. *Journal of chemical ecology.* 2002; 28:2371–2376. [PubMed: 12523574]
5. Xu PX, Zwiebel LJ, Smith DP. Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. *Insect Mol Biol.* 2003; 12:549–560. [PubMed: 14986916]
6. Biessmann H, Nguyen QK, Le D, Walter MF. Microarray-based survey of a subset of putative olfactory genes in the mosquito *Anopheles gambiae*. *Insect Mol Biol.* 2005; 14:575–589. [PubMed: 16313558]
7. Foret S, Maleszka R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* 2006; 16:1404–1413. [PubMed: 17065610]
8. Wogulis M, Morgan T, Ishida Y, Leal WS, Wilson DK. The crystal structure of an odorant binding protein from *Anopheles gambiae*: evidence for a common ligand release mechanism. *Biochem Biophys Res Commun.* 2006; 339(1):157–164. [PubMed: 16300742]
9. Tsitsanou KE, Thireou T, Drakou CE, Koussis K, Keramioti MV, Leonidas DD, Eliopoulos E, Iatrou K, Zographos SE. *Anopheles gambiae* odorant binding protein crystal complex with the synthetic repellent DEET: implications for structure-based design of novel mosquito repellents. *Cell Mol Life Sci.* 2011 Jun 14. [Epub ahead of print].
10. Biessmann H, Andronopoulou E, Biessmann MR, Douris V, Dimitratos SD, Eliopoulos E, Guerin P, Iatrou K, Justice RW, Kröber T, Marinotti O, Woods DF, Walter MF. The *Anopheles gambiae* Odorant Binding Protein 1 (angaOBP1) mediates indole recognition in the antennae of female mosquitoes. *PLoS One.* 2010; 5(3):e9741. [PubMed: 20305819]
11. Mao Y, Xu X, Xu W, Ishida Y, Leal WS, Ames JB, Clardy J. Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. *Proc Natl Acad Sci U S A.* 2010; 107(44):19102–19107. [PubMed: 20956299]
12. Lartigue A, Gruez A, Briand L, Blon F, Bezirard V, Walsh M, Pernollet JC, Tegoni M, Cambillau C. Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. *The Journal of biological chemistry.* 2004; 279:4459–4464. [PubMed: 14594955]
13. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *J. Computational Chemistry.* 1998; 19:1639–1662.
14. Poupon A, Mornon JP. "Topohydrophobic positions" as key markers of globular protein folds. *Theor Chem Accounts.* 1999; 101:2–8.
15. Poupon A, Mornon JP. Populations of hydrophobic amino acids within protein globular domains: identification of conserved "topohydrophobic" positions. *Proteins.* 1998; 33:329–342. [PubMed: 9829693]
16. Poupon A, Mornon JP. Predicting the protein folding nucleus from sequences. *FEBS Lett.* 1999; 452:283–289. [PubMed: 10386607]
17. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 1995; 247:536–540. [PubMed: 7723011]
18. Andreeva A, Howorth D, Chandonia JM, Brenner SE, Hubbard TJ, Chothia C, Murzin AG. Data growth and its impact on the SCOP database: new developments. *Nucl. Acids Res.* 2008; 36:D419–D425. [PubMed: 18000004]

19. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucl. Acids Res. 2010; 38:W545–W549. [PubMed: 20457744]
20. Berezovsky IN, Grosberg AY, Trifonov EN. Closed loops of nearly standard size: common basic element of protein structure. FEBS Lett. 2000; 466:283–286. [PubMed: 10682844]
21. Berezovsky IN, Kirzhner VM, Kirzhner A, Trifonov EN. Protein folding: looping from hydrophobic nuclei. Proteins. 2001; 45:346–350. [PubMed: 11746682]
22. Lamarine M, Mornon JP, Berezovsky IN, Chomilier J. Distribution of tightened end fragments of globular proteins statistically matches that of topohydrophobic positions: towards an efficient punctuation of protein folding? Cell. Mol. Life Sci. 2001; 58:492–498. [PubMed: 11315195]
23. Rehm BH, Qi Q, Beermann BB, Hinz HJ, Steinbuechel A. Matrix-assisted in vitro refolding of *Pseudomonas aeruginosa* class II polyhydroxyalkanoate synthase from inclusion bodies produced in recombinant *Escherichia coli*. Biochem. J. 2001; 358:263–268. [PubMed: 11485576]
24. Provencher SW, Glockner J. Estimation of globular protein secondary structure from circular dichroism. Biochemistry. 1981; 20:33–37. [PubMed: 7470476]
25. Sreerama N, Woody RW. Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. J. Mol. Biol. 1994; 242:497–507. [PubMed: 7932706]
26. Qiao H, He X, Schymura D, Ban L, Field L, Dani FR, Michelucci E, Caputo B, della Torre A, Iatrou K, Zhou JJ, Krieger J, Pelosi P. Cooperative interactions between odorant-binding proteins of *Anopheles gambiae*. Cell Mol Life Sci. 2011; 68(10):799–813.
27. Bhonsle JB, Bhattacharjee AK, Gupta RK. Novel semi-automated methodology for developing highly predictive QSAR models: application for development of QSAR models for insect repellent amides. J Mol Model. 2007; 13(1):179–208. [PubMed: 17048015]
28. Browne JP, Strom M, Martin SR, Bayley PM. The role of beta-sheet interactions in domain stability, folding, and target recognition reactions of calmodulin. Biochemistry. 1997; 36:9550–9561. [PubMed: 9236001]
29. Fasman, GD. Differentiation between transmembrane helices and peripheral helices by deconvolution of circular dichroism spectra of membrane proteins. In: Fasman, GD., editor. Circular Dichroism and the Conformational Analysis of Biomolecules. New York: Plenum Press; 1996. p. 381–412.
30. Brahms S, Brahms J. Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. J. Mol. Biol. 1980; 138:149–178. [PubMed: 7411608]
31. Venyaminov, SY.; Yang, JT. Determination of protein secondary structure. In: Fasman, GD., editor. Circular Dichroism and the Conformational Analysis of Biomolecules. New York: Plenum Press; 1996. p. 69–107.
32. Woody, RW.; Dunker, AK. Aromatic and cystine side-chain circular dichroism in proteins. In: Fasman, GD., editor. Circular Dichroism and the Conformational Analysis of Biomolecules. New York: Plenum Press; 1996. p. 109–157.
33. Martin SR, Schilstra MJ. Circular dichroism and its application to the study of biomolecules. Methods Cell Biol. 2008; 84:263–293. [PubMed: 17964935]
34. DeLano, WL. The PyMOL Molecular Graphics System on World Wide Web. 2004. <http://www.pymol.org> Thireou T, Atlamazoglou V, Papandreou NA, Lonquety M, Chomilier J, Eliopoulos E. Quantitative prediction of critical amino acid positions for protein folding. Protein Pept Lett. 2009; 16:1342–1349. [PubMed: 19508208]
35. Whitmore L, Wallace BA. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. Biopolymers. 2008; 89:392–400. [PubMed: 17896349]

Highlights

- *In silico* protein structure modeling and site-directed mutagenesis.
- Circular dichroism and ligand-binding assays to probe the effects of the point mutations.
- Interpretation of mutational effects on folding and ligand binding.

Figure 1.

In silico analysis of *Anopheles gambiae* OBP1. (A) TH positions and TEFs along the aminoacid sequence of AgamOBP1 together with the performed mutations. Line A indicates the secondary structure elements (regular α -helix 66%, regular β -strand 2%, turns, disordered and non regular elements 32%) together with the disulfide bridges, line B the primary structure and the TH positions and line C the mutations performed. (B) OBP1 main chain is shown in schematic ribbon representation in grey with TH and TEF positions colored green. Aminoacid side chains in cyan are those of residues mutated and producing soluble protein. Side chains in pink are those of residues mutated and producing insoluble protein. (C) Low energy conformation of 1-NPN in the binding pocket. Diagrams created with PYMOL [33] (DeLano 2004).

Figure 2. Native OBPI (in the center) and point mutant proteins (1. M84Q 2. L15T 3. L19T 4. M84K 5. Y54A 6. L58T 7. F59Y 8. W114Q). Cartoon representations are colored based on the Folding Profile [34]. Colors range from blue to red, corresponding to the minimum and maximum mean number of spatial neighbors respectively. Amino acid point mutations change the OBPI Folding Profile reflecting changes in protein conformation, in the percentage and orientation of secondary structure elements.

Figure 3.

2 Circular Dichroism spectra of r*Agam*OBP1 and mutated r*Agam*OBP1 Far-UV CD spectrum of wild type *Agam*OBP1 shows a maximum at 192 nm and two minima at 208 and 225 nm, typical of a protein with a significant α -helical secondary structure and is integrated in all four panels. Exchange of polar tyrosine for non-polar phenylalanine at residue 123 (panel I). Substitution of polar glutamine for nonpolar methionine at residue 84 (panel II). *Agam*OBP1 mutations resulting in major losses of α -helical structure (panel III). Mutations resulting in maximum loss of α -helical structure and shift of the 208nm minimum (panel IV).

Figure 4.

Binding of 1-NPN to rAgamOBP1 and rAgamOBP1 mutants. 1-NPN was tested at 8 μ M and the concentration of recombinant proteins was 4 μ M (higher curves in duplicate). A. Binding curves for native AgamOBP1. B. Binding curves for mutant F123T. C. Binding curves for mutant H111A. Addition of 16 μ M 3-methyl indole to the assays results in the displacement of 1-NPN from the protein binding pocket and a reduction in fluorescence (lower curves in duplicate).

Table 1

List of rAgamOBP1 and rAgamOBP1 mutants expressed, purified and analyzed in this study. Proteins accumulated in inclusion bodies after expression in *E. coli*, labeled “Insoluble”, were denatured and refolded resulting in soluble proteins. The CD spectra were utilized for the determinations of CD spectra and 1-NPN binding properties. The CD spectra were classified in four types (I,II,III and IV), being that type I is the most similar to the data from rAgamOBP1 and type IV the most divergent. The areas under the fluorescence curves were used for calculating relative 1-NPN binding efficiency for the different proteins. The percentage value was calculated as the area under data collected from 400 nm to 440 nm and exceeding 200 RU (response units), with the maximum observed for rAgamOBP1 set as 100%.

	Mutations		Change	Soluble/ Insoluble	CD spectrum type	1-NPN binding
1	rAgamOBP1	-----	-----	Soluble	I	100%
2	L58T	Leucine to Threonine	Nonpolar to Polar	Insoluble refolded	IV	85%
3	M84Q	Methionine to Glutamine	Nonpolar to Polar	Soluble	II	78%
4	Y122F	Tyrosine to Phenylalanine	Polar to Nonpolar	Soluble	III	75%
5	F123Y	Phenylalanine to Tyrosine	Nonpolar to Polar	Soluble	I	71%
6	M84K	Methionine to Lysine	Nonpolar to Polar	Insoluble refolded	III	69%
7	F123T	Phenylalanine to Threonine	Nonpolar to Polar	Insoluble refolded	III	64%
8	H111A	Histidine to Alanine	Polar to Nonpolar	Soluble	III	64%
9	Y54F	Tyrosine to Phenylalanine	Polar to Nonpolar	Insoluble refolded	III	63%
10	F59Y	Phenylalanine to Tyrosine	Nonpolar to Polar	Insoluble refolded	IV	61%
11	L19T	Leucine to Threonine	Nonpolar to Polar	Insoluble refolded	III	59%
12	Y122A	Tyrosine to Alanine	Polar to Nonpolar	Insoluble refolded	III	56%
13	Y54A	Tyrosine to Alanine	Polar to Nonpolar	Insoluble refolded	IV	52%
14	L38Y	Leucine to Tyrosine	Nonpolar to Polar	Insoluble refolded	III	51%
15	L15T	Leucine to Threonine	Nonpolar to Polar	Soluble	III	49%

	Mutations		Change	Soluble/ Insoluble	CD spectrum type	1-NPN binding
16	W114Q	Tryptophan to Glutamine	Nonpolar to Polar	Soluble	IV	46%
17	L124T	Leucine to Threonine	Nonpolar to Polar	Insoluble refolded	III	46%

Table 2

Secondary structure of recombinant AgamOBP1 mutant proteins calculated from their far-UV CD spectra. Helix1 represents calculated regular α -helix structure. Helix2 represents distorted α -helix structure. Δ Helix1 and Δ Helix2 represent the change in calculated α -helix structure from the native AgamOBP1 protein resulting from the specific amino acid substitution. Secondary structure changes were calculated from the un-normalized CD data in the far-UV region using algorithms available on the DicroWeb website [35]. Several deconvolution algorithms and appropriate reference sets were evaluated. The CONTIN deconvolution algorithm was chosen based on the range of input data obtained. Reference sets contained CD spectra of globular proteins of known X-ray crystal structure, representative of α -helix, β -sheet, β -turn and unordered elements [25]. The best reference set was selected on the basis of the NRMDS (normalized root mean square deviation) value obtained when the data were analyzed with various reference sets. Reference sets 4 and 7, which contain 43 and 48 proteins respectively, gave the most accurate results and the lowest NRMDS values. Although the deconvoluted data obtained from circular dichroism does not duplicate the secondary structure percentage assignment obtained from X-ray crystallography, each of the CD scans of the OBP1 Native and mutant proteins were deconvoluted using the CONTIN algorithm with protein reference set #7 that gave the best fit and lowest NRMDS [25]. Therefore, there is an internal consistency in the changes calculated for the mutants ($\Delta\alpha$ -helix, $\Delta\beta$ -strand, Δ turn, Δ unordered).

Protein	NRM SD	Helix 1	Helix 2	Strand 1	Strand 2	Turns	Unord	Δ Helix 1	Δ Helix 2	Δ Strand 1	Δ Strand 2	Δ Turnl	Δ Unord	Helix Sum
OBP1 Native	0,091	65,9	32,8	0	1,3	0	0							98,7
L15 T	0,131	47,5	39,3	0	1,1	2,4	9,7	-18,4	6,5	0,0	-0,2	2,4	9,7	86,8
M84 Q	0,211	58,7	41,3	0	0	0	0	-7,2	8,5	0,0	-1,3	0,0	0	100,0
H11 LA	0,171	55,3	44,7	0	0	0	0	-10,6	11,9	0,0	-1,3	0,0	0	100,0
W11 4Q	0,147	47,8	40,3	0	0	3,2	8,8	-18,1	7,5	0,0	-1,3	3,2	8,8	88,1
Y122 F	0,090	43,2	31	0	0	6,2	19,5	-22,7	-1,8	0,0	-1,3	6,2	19,5	74,2
F123 Y	0,074	63,1	36,6	0	0,3	0	0	-2,8	3,8	0,0	-1,0	0,0	0	99,7
L19 T	0,144	42,9	56,5	0	0,6	0	0	-23,0	23,7	0,0	-0,7	0,0	0	99,4
L38 Y	0,091	44,5	30,9	0	0	6,5	18	-21,4	-1,9	0,0	-1,3	6,5	18	75,4
Y54 A	0,202	62,1	37,9	0	0	0	0	-3,8	5,1	0,0	-1,3	0,0	0	100,0
Y54 F	0,167	56,1	43,9	0	0	0	0	-9,8	11,1	0,0	-1,3	0,0	0	100,0
L58 T	0,166	51,4	41,3	0	0	7,3	0	-14,5	8,5	0,0	-1,3	7,3	0	92,7

Protein	NRM SD	Helix 1	Helix 2	Strand 1	Strand 2	Turns	Unord	Δ Helix 1	Δ Helix 2	Δ Strand 1	Δ Strand 2	Δ Turnl	Δ Unord	Helix Sum
F59 Y	0,149	47, 6	40, 1	0	0	5,5	6,8	-18,3	7,3	0,0	-1,3	5,5	6,8	87, 7
M84 K	0,154	57, 1	41, 4	0	0	1,6	0	-8,8	8,6	0,0	-1,3	1,6	0	98, 5
Y122 A	0,097	47	34, 2	0	0	6,2	12,7	-18,9	1,4	0,0	-1,3	6,2	12,7	81, 2
F123 T	0,091	44, 5	30, 8	0	0	6,4	18,3	-21,4	-2,0	0,0	-1,3	6,4	18,3	75, 3
L124 T	0,098	49, 3	37, 1	0	0	5,7	7,9	-16,6	4,3	0,0	-1,3	5,7	7,9	86, 4