Nuclear Magnetic Resonance Studies of Allergens

Paul Rösch, Dept. of Biopolymers, University of Bayreuth, 95440 Bayreuth, Germany

Introduction

A variety of allergens is known that are recognized by IgE and IgG antibodies either on the basis of linear epitopes or on the basis of conformational epitopes [1]. From a structural biology point of view, protein and peptide antigens eliciting a B−cell response may thus be divided into two major categories, namely those that contain an uninterrupted sequence of amino acids recognized by antibodies and those that are recognized on the basis of their spacial structure, that is the three−dimensional arrangement of sequences of key amino acids. The latter type of allergens include the major birch pollen allergen Bet v 1 [2,3], the cherry allergen Pru av 1 [4,5], the dust mite allergens Der p 2 and Der f 2 [6,7], profilin [8,9], the cat allergen Fel d 1 [10,11], and the major giant ragweed allergens Amb a 5 and Amb t 5 [12,13].

Understanding of protein action on a molecular level was mainly determined by the progress of biophysical techniques such as X−ray crystallography and nuclear magnetic resonance (NMR) as applied to studies of biopolymers. Whereas X−ray crystallography was established as a structure determination tool for proteins and nucleic acids more than forty years ago, the application of NMR as a technique in structural molecular biology is more recent, and it was only about 20 years ago that the first protein structures emerged from NMR data [14], although NMR was used in the general study of proteins and nucleic acids and their interactions much earlier. Indeed, in recent years the most spectacular results were achieved by use of NMR as a spectroscopic tool that, in combination with molecular dynamics calculations, allowed the determination of complete three−dimensional structures of biomolecules with a quality rivaling that of structures determined by X−ray crystallography. NMR spectroscopy is, however, not restricted to the determination of molecular structures in atomic detail as such, but allows study of a broad range of molecular phenomena in the biological sciences, such as description of the time dependence of molecular structures, determination of reaction and binding kinetics, and others [15,16].

The isolation of allergens from plant and animal material and the subsequent cloning, expression, and purification procedures that resulted in the availability of highly purified material were the prerequisites for the application of advanced multidimensional NMR−techniques to study the molecular structure of these molecules in atomic detail. NMR structure determination procedures are traditionally thought to be restricted to low molecular weight molecules, such as proteins below a molecular mass of about 25,000, but more recent techniques hold strong promise that biomolecules of considerably higher mass can be subjected to NMR studies in the not so distant future [17, 18]. Many allergens that are defined on a molecular level have a molecular mass below the 25,000 limit and are thus accessible to current NMR structure determination techniques, for example the cherry allergen Pru av 1, the birch pollen allergens Bet v 1 and Bet v 4, the timothy grass pollen allergen Phl p 2, the house dust mite allergen Der p 2, the giant ragweed allergen Amb t 5, the peanut allergen Ara h 2, and others. To date, several of these allergens resisted all attempts of crystallization, and advantage was taken of the unique capability of NMR to determine the three−dimensional structure of biomolecules in atomic resolution in solution to determine their conformation.

The conformations of several allergens have been determined by classical two−dimensional or heteronuclear higher−dimensional NMR methods that required isotopic labeling with ^{15}N and ^{13}C (table 1). Whereas Amb a 5, and Amb t 5 are small peptides of 45 and 40 amino acids, respectively, Der p 2, Bet v 1, Bet v 4, and Pru av 1, are proteins of 85 to 159 amino acids. Accordingly, the structure of the smaller allergens could be solved using exclusively two−dimensional techniques, but the structure determination procedures for Bet v 1, Bet v 4, and Pru av 1 had to resort to higher−dimensional techniques with the prerequisite of the availability of fully labeled material. The general topology of the protein allergens known in conformational detail so far varies widely, ranging from EF−hand type calcium ion binding structures to $\alpha-\beta$ proteins and β−barrels.

In spite of all the current structural information brought about by NMR and X−ray crystallography, the biological function of most protein allergens is still in the dark. To sufferers from allergies it may be an academic question, but for fundamental reasons and in the believe that nature did not invent protein allergens for the sole purpose of annoying mankind, the search for the physiological function

of these proteins goes on. A particular help is the comparison of the three−dimensional fold of the protein allergens with the database of proteins whose topology and biological function is known so as to conclude from a similar structure on a similar function.

NMR and protein solution structures

The determination of solution structures of proteins as it is most widely used is divided into three essential parts [15,16]: sequence−specific assignment of resonances in the NMR−spectra, collection of NMR data to define approximate distances of protons, and molecular dynamics calculations that are designed to search conformational space for three−dimensional structures compatible with the experimentally determined distances. The NMR experiments rely on two fundamentally different interactions of nuclei, one of them is mediated by chemical bonds, that is it is a through−bond interaction (J−coupling), the other one is a through space interaction (dipolar coupling) [19, 20].

The first major step in protein structure determination by NMR is the assignment of resonances in the spectra to specific protons in the molecule. In lower molecular weight systems that are well behaved, this assignment procedure can mostly be performed with the help of standard two−dimensional experiments that reveal the chemical bond connectivities between protons via their J−coupling correlation, and the short−range spacial connectivities by their dipolar coupling. Dealing with protein molecules of higher molecular weight demands the introduction of higher−dimensional, heteronuclear NMR−methods, usually in three−dimensional techniques, for the assignment procedures. These higher−dimensional experiments also reveal J−coupling connectivities, but these J−couplings are now detected not only between protons, but between protons and other magnetically active nuclei, for example between protons, ^{13}C−, and ^{15}N−nuclei. As the natural abundance of ^{15}N− and ^{13}C− is too low for these higher−dimensional techniques to be applied, availability of fully isotope labeled material is necessary. Thus, in most cases, bacterial expression systems with very high yield (tens of milligrams per liter culture) are mandatory, as the cost of the isotope labeled culture media (^{13}C−glucose and ^{15}N−ammonium sulfate are used most widely) would otherwise be prohibitive.

The second step in the solution structure determination of proteins by NMR is the measurement of distances between protons. These measurements rely on the through space dipolar coupling which is experimentally manifested in the nuclear Overhauser effect (NOE). The NOEs are detected as crosspeaks in standard two or higher dimensional NMR experiments. Due to the dependence of the intensity of the NOE−effect on r^{-6}, where r is the distance between the protons involved, the dipolar coupling is a very short range interaction, and an NOE between protons is generally assumed to be measurable only if the distance between these protons is less than 0.6 nm. As the NOE is also influenced by time variations of the dipolar interactions and thus by the dynamics of the protein molecules, distances cannot be determined with high accuracy [20]. In addition to the NOE data, other parameters that define spatial relations between atoms such as the bond−angle dependent J−couplings or, more recently, the dependence of the dipolar coupling on the angle between the distance vector of two coupled protons and the external magnetic field may be used [21].

The third step is the search for protein conformations that are compatible with the experimental data. This is usually done by a molecular dynamics calculation involving a so−called simulated annealing procedure to overcome local energetic minima [22]. To this end, the distance and angle constraints are converted to energy functions with minima corresponding to the approximate values that were determined experimentally. The result of these calculations is then a family of structures all of which fulfill the experimental data. The variation of the structures within these families is measured as root mean square deviation and depends, among others, on the flexibility of the molecules and the number and accuracy of the experimental constraints.

Solution Structures of individual allergens

Ragweed pollen allergens Amb a 5 and Amb t 5

Ragweeds *(Ambrosia)* contain peptide allergens of about 40 amino acids, and interest focused early on the two Amb allergens Amb t 5 *(Ambrosia trifida)* and Amb a 5 *(Ambrosia artemisiifolia).* These allergens are 46.3% identical in their amino acid sequence in a 41 residue overlap [12,13] (table I). They possess similar MHC II binding regions [23], but their T cell receptor interacting residues are different [24], and they are not cross−reactive on the antibody level [25]. Such comparatively high sequence identity is not necessarily a safe indication of high identity on the level of tertiary structure, as recent examples clearly show [26]. As a good initial working hypothesis, however, the assumption that nearly 50 % sequence identity results in highly similar three−dimensional structures may be accepted. This observation suggest that study of the conformations of Amb t 5 and Amb a 5 in atomic detail could further our understanding of the causes of cross allergies and provide some clues as to the nature of the conformational epitopes of these peptides. This, in combination with the fact that peptides of the size of the Amb a 5 and Amb t 5 allergens are easily accessible to structure determination by conventional NMR−techniques, led to efforts to study the structures of the wildtypes of these allergens in solution and combine the solution structural studies of the wildtype peptides with studies of the solution conformation and the allergenic potential of designed mutant peptides. Consequently, Amb a 5 and Amb t 5 were among the first allergens for which NMR spectra were available (Amb a 5: [27]) and whose structures could be determined in atomic resolution by NMR [28,29,30,31].

Demand of material for structural studies is usually in the milligram range. Thus, both allergens were expressed in *E. coli*, and antibody binding studies of the recombinant material showed recombinant Amb t 5 to be identical to the natural peptide, a claim supported by the exact agreement of their NMR spectra [28], This was not self−evident as the Amb t 5 peptide contains four disulfide bonds. In fact, recombinant Amb a 5 was not identical to the natural product, making purification from natural sources necessary for the structure determination procedures [29]. As the peptides are well−structured and their size is modest, it was not necessary to resort to multidimensional NMR−techniques with isotope labeled material, and the structure determination could be performed by standard homonuclear techniques [15] and molecular dynamics calculations [22]. The proton resonance assignments were essentially complete for both, with exception of Gly8 in Amb t 5.

The Amb a 5 and Amb t 5 structures are highly identical as far as their general topology is concerned. According to the entry in the protein data bank [32], Amb t 5 shows a COOH−terminal helix that extends from Lys32 to Asn36, and two antiparallel β−strands that run from Tyr17 to Cys19 and from Val27 to Tyr29 (fig. 1). More extended secondary structure elements and the presence of a third short β−strand, from Cys5 to Glu7, was reported in the original publication [29]. During the course of data refinement, the correct pairing of disulfide bonds in Amb t 5 as Cys5−Cys35, Cys11−Cys26, Cys18−Cys28, and Cys19−Cys39 became apparent. The final conformation was confirmed by the fact that several extraordinary chemical shifts could be explained nicely by the structure data. The Amb a 5 structure [30], also determined by standard two−dimensional NMR and molecular dynamics methods, shows the same overall topology as the Amb t 5 structure, and a secondary structure was reported consisting of a COOH−terminal helix from 35−43 and three β−strands, from Cys4−Ala7, from Tyr17−Ser20, and fromVal31−Tyr33. The β−strand not included in the pdb entry for Amb t 5 is less well defined than the other two β−strands according to the local root mean square deviation (rmsd) data [29].

As a consequence of these studies, variants of Amb t 5 were constructed in an attempt to define the B−cell epitope. Amino acids were exchanged for the homologues residues of Amb a 5, with the plan to alter the immunogenic sites of Amb t 5 to those of Amb a 5, a goal that could not be achieved completely as properties of recombinant Amb a 5 did not match those of the natural product, and thus replacement of too many residues seemed inappropriate. The epitope, however, could be narrowed to amino acids in loop 3 of the structure (fig. 1) [30].

Dust mite allergens Der f 2 and Der p 2

The 129 amino acid dust mite *(Dermatophagoides pteronyssinus)* major allergens Der f 2 and Der p 2 are identical in sequence to 87.5 % in a 128 amino acid overlap and are thus expected to be virtually identical on the level of tertiary structure [6,7] (table I). Both proteins exhibit three disulfide bonds whose presence was judged to be important for IgE binding [33,34]. In Der p 2, these disulfide bonds connect cysteines 8 and 119, 21 and 27, and 73 and 78.

Structure determination of the dust mite allergens Der f 2 and Der p 2 necessitated use of isotope labeled protein. As a prerequisite, Der f 2 as well as Der p 2 were expressed in *E. coli* systems efficient enough to allow ^{13}C and ^{15}N isotope labeling. Three−dimensional triple resonance NMR−techniques and molecular dynamics calculations were then applied to obtain the complete conformations of Der f 2 [35] and Der p 2 [36] (fig. 2). The structure determinations were performed under different conditions, and more extensive NMR−experiments as well as double−labeling of specific amino acids resulted in a clearly higher number of observable long range NOEs and a higher quality of the Der p 2 structure [36]. These allergens are predominantly β proteins, with Der p 2 exhibiting six β−strands, from K15 to L17, F35 to E42, K51 to I58, Q85 to W92, V104to M111, and L117 to A122, and two weak, short helices, from I58 to G60 and A72to H74. The weakness of the helices allows classification of the protein as all−β. Comparison of either protein tertiary structure with the database using the DALI program [37] yielded classification of the fold as a six stranded immunoglobulin type sandwich arrangement of β−strands with a topology reflecting the s−type as defined earlier [38]. This fold, however, is extremely common in nature, and thus does not provide much clue to the physiological function of the protein. A COOH−terminal α−helix that was proposed to be a general major contributor to the allergenicity of proteins [39] is not present in either Der p 2 or Der f 2. At the time of this writing, a similarity search [37] returned the structure of the rho guanine nucleotide dissociation inhibitor as the closest structural relative. Whether or not this resemblance with a z−score of 5.9 indicates any functional similarity remains to be seen. From a comparison of amino acid sequences, a function of Der p 2 in molting has been suggested [36].

Timothy grass pollen allergen Phl p 2

The 96 amino acid timothy grass pollen allergen Phl p 2 was expressed in *E. coli* and shown to be equivalent to the natural protein [40]. The two−dimensional NMR spectra were of excellent quality, and resonances were extremely well resolved, so that the proton resonance assignment could be performed without resorting to isotope labeled protein and multidimensional experiments [41]. The assignment was complete with the exception of loop region Leu60−Pro64. Phl p 2 turned out to be a protein of the all−β type, with nine strands as identified by long range HN−HN, HN−Nα Hα−Hα, HN−Hβ NOE connectivities. According to analysis with the DALI program [37], the structure of Phl p 2 (fig. 3) resembles an immunoglobin fold (z−scores to related immunoglobulin−type proteins around 4.9) topologically related to the h−type [38]. The first β−sheet of the Phl p 2 immunoglobulin fold is formed by strands running from 42−44, 29−35, 64−70, and 74−80, the second β−sheet is formed by strands 6−9, 16−21,.45−47, 51−56, and 90−93. The resemblance of Phl p 2 to immunoglobulin folds, however, is not as close as that of Der p 2 (z−scores around 5.6]. In addition, the architecture of Phl p 2 does not show the core residues that are generally assumed defining the immunoglobulin fold family of proteins [38]. Like the Der f 2 and the Der p 2 three−dimensional structure, the Phl p 2 structure does not show any sign of the presence of an α−helical secondary structure.

Birch pollen allergen Bet v 1 and cherry allergen Pru av 1

The largest protein allergens whose structures have been determined so far by high−resolution NMR−spectroscopy are the major birch pollen allergen Bet v 1 and the cherry allergen Pru av 1, both of which consist of 159 amino acids. Bet v 1 and Pru av 1 are 60 % identical in their sequences (table I). Initially, the COOH−terminal sequence region was shown to form an α−helix independent of the protein structural context as determined by two−dimensional NMR−spectroscopy on an isolated peptide [42]. This helix is amphipathic. Bet v 1 was judged to be a monomeric protein from NMR and circular dichroism studies [43]. The secondary structure and the global fold of Bet v 1 were elucidated by NMR [42], and subsequently the three−dimensional structure could be determined by X−ray crystallography as well as multidimensional triple resonance NMR using ^{15}N labeled and ^{15}N−^{13}C double−labeled material [44,45]. The same methods were used to determine the three−dimensional structure of Pru av 1 [46,47] (fig. 4). The structures of Pru av 1 and Bet v 1 turned out to be virtually identical, with an rmsd value of less than 0.2 nm for the solution structure of Pru av 1 vs. the crystal structure of Bet v 1 [47] (fig. 5). This rmsd is roughly the same as the deviation of the crystal and the NMR−structures (0.16 nm; [44]]. The conformation of Bet v 1 features a large β−sheet consisting of seven β−strands, from Thr7 to Ser11, Val41 to Glu45, Ile53 to Ile56, Arg70 to Asp75, Lys80 to Ser84, Gly111 to His121, Lys97 to Thr107, and three α−helices, from Ala15 to Phe22, Gly26 to Val33, and Ala130 to Ala153, resulting in a βααββββββα–topology. The arrangement of β−strands is distantly resembling the »β−cup« arrangement found in the lipocalin protein family, although the latter motif consists typically of an 8−stranded partial barrel that is rather closed, with a COOH−terminal α−helix on top of the barrel [48]. In the context of the whole protein, the hydrophobic residues of the COOH−terminal helix as well as the hydrophobic residues of the two other helices are pointing to a large basket formed by the β−sheets. The three helices are in a nearly perpendicular arrangement to each other with the two short helices α1 and α2 under the long COOH−terminal helix α3. This helix arrangement and the position of the α3−helix is stabilized by a hydrophobic cluster consisting of Leu18, Ala21, Phe22, Leu29, Val147. The β−sheets thus form a basket, with α3 as a lid and α1 and α2 forming a clip (fig. 6). This arrangement stabilizes the orientation of the single methionine residue, Met139, and of residue Leu143 pointing into the center of the cavity formed by the COOH−terminal helix, helices α1 and α2, and the β−sheet. The loop following β7, Thr122to Lys129, forms the hinge of the α3−lid.

Until recently, the only hint towards the function of these proteins came from a report of RNAse activity of Bet v 1 [49]. Comparison of the Bet v 1 (pdb−entry: 1B6F) or the Pru av 1 (pdb−entry: 1E09) NMR−structure with the pdb database, however, resulted in a single, very close match (z−score above 10) with a recently determined structure of the human protein mln64 ([50]; pdb−entry: 1EM2). This steroidogenic acute regulatory protein (StAR) regulates acute steroidogenesis in the adrenal cortex and gonads by promoting the translocation of cholesterol to the mitochondrial inner membrane where the first step in steroid biosynthesis is catalyzed. The Bet v 1 and Pru av 1 allergens show only very low sequence homology with mln64. Like Bet v 1 and Pru av 1, mln64 shows an architecture of an extended β−sheet and α−helices surrounding a large cavity assumed to be the cholesterol binding cavity which, by analogy, may be inferred to be a plant steroid binding site in the Bet v 1/Pru av 1 type plant allergens. Surprisingly, the allergen and the mln64 structures agree even in a minute, but remarkable detail, namely the lid−clip formed by the COOH−terminal helix α3 and the two small helices α1 and α2. In mln64, the arrangement is stabilized by a hydrophobic core of Leu295, Val296, Val300, Leu434, and Ile438. Two of the three methionine residues, Met307 and Met388, are very close to the lid−clip arrangement, and the orientation of Met427 is fixed by the helix−basket motif in a direction very similar to Met139 of Bet v 1. From this similarity, it would not be surprising if the Bet v 1/Pru av 1 family of allergens would consist of plant steroid binding proteins involved, for example, in signal transduction.

Indeed, recent NMR−experiments with Pru av 1 and the plant hormone homocastasterone indicated that plant steroids bind to the hydrophobic cavity of this protein, and molecular modeling calculations showed that two castasterone molecules can simultaneously occupy the cavities of Bet v 1 or Pru av

1 [51] (fig. 7). The surprisingly high similarity of the structures of the cherry and the birch pollen allergens has been suggested to partially explain the cross−allergies observed for these allergens [51].

Outlook

One the major aspects of allergenic activity of proteins that needs still to be studied and may profit tremendously by the application of NMR−methods is the interaction of allergens and antibodies and the search for structural epitopes. With the limited data currently accessible from NMR−studies of protein allergens (table II), however, the search for epitopes is mainly restricted to the comparison of structural features of mutants, as has been done with the Amb 5 allergens, or comparison of highly similar proteins, for example Amb a 5 / Amb t 5, Der f 2 / Der p 2, and Bet v 1 / Pru av 1, and, as of present, no generalization of any distinctive structural element that may determine the allergenic properties of these compounds could be derived. In particular, only rarely have the specific advantages of NMR been used in the study of allergens, that is the characterization of dynamic properties of allergens or the determination of surface contacts between allergens and antibodies. In particular the latter field of studies, that is, in practical terms, the determination of structural features of allergen−F_{ab}−fragment complexes or at least single−chain F_{ab}−fragments is highly likely to become a major focus of NMR studies in the near future. Several experiments towards this end were performed already, albeit not directly aimed at allergen−antibody complexes. Definition of interfaces of large protein−protein complexes in solution by high resolution NMR has been tackled by use of experiments involving chemical shift perturbation, isotope exchange techniques involving main chain amide groups, or specific labeling of amino acids. For example, a structure could be suggested for the HIV−1 anti−gp120 antibody complex with a small peptide by extensive use of deuterium labeling techniques [52], and the exploitation of cross−saturation phenomena detected in $^{1}H-^{15}N$ heteronuclear single quantum coherence experiments with TROSY coherence transfer in an optimally deuterium labeled system was demonstrated to yield valuable information on large (M_r 64,000) protein−protein complexes [53]. Thus, although X−ray crystallography may be the superior technique for the study of larger protein complexes, the current advances in NMR−technology as well as the increasing experience with the appropriate handling of these complexes makes them accessible to NMR−studies, thus opening up all the advantages of this technique such as its power in studying the dynamics and kinetics of interactions or the possibility to look at interaction surfaces in detail in solution.

Acknowledgement

This article was written during a sabbatical stay in Dr. Peter Wright’s group at The Scripps Research Institute. I would like to thank him and all group members for the pleasant atmosphere and the Deutsche Forschungsgemeinschaft (Ro617/12−1) for making this stay possible.

Literature

[1] M.H. van Regenmortel, J.L. Pellequer, Pept. Res. 7 (1994) 224

[2] H. Ipsen, H. Loewenstein J. Allergy Clin. Immunol. 72 (1983) 150

[3] E. Jensen−Jarolim, A. Leitner, H. Kalchhauser, A. Zurcher, E. Ganglberger, B. Bohle, O. Scheiner, G. Boltz−Nitulescu, H. Breiteneder, FASEB J. 12 (1998) 1635

[4] S. Schöning, S. Vieths, A. Petersen, W. Baltes, J. Sci. Food Agric. 67 (1995) 431

[5] S. Scheurer, D.Y. Son, M. Boehm, F. Karamloo, S. Franke, A. Hoffmann, D. Haustein, S. Vieths, Mol Immunol. 36 (1999) 155

[6] P.W. Heymann, M.D. Chapman, R.C. Aalberse, J.W. Fox, T.A.E. Platts−Mills, J. Allergy Clin. Immunol. 83 (1989) 1055

[7] C. Nishiyama, M. Fukada, Y. Usui, N. Iwamoto, T. Yuuki, Y. Okumura, H. Okudaira, Mol. Immunol. 32 (1995) 1021

[8] A. Leitner, M. Vogel, C. Radauer, H. Breiteneder, B.M. Stadler, O. Scheiner, D. Kraft, E. Jensen−Jarolim, Eur. J. Immunol. 28 (1998) 2921

[9] R.H. Sohn, P.J. Goldschmidt−Clermont, Bioessays 16 (1994) 465

[10] I.J. Griffith, S. Craig, J. Pollock, X.B. Yu, J.P. Morgenstern, B.L. Rogers, Gene 113 (1992) 263

[11] W. van ’t Hof, F.J. van Milligen, M. van den Berg, M. Lombardero, M.D. Chapman, R.C. Aalberse, Allergy 48 (1993) 255

[12] L.E. Mole, L. Goodfriend, C.B. Lapkoff, J.M. Kehoe, J.D. Capra, Biochemistry 14 (1975) 1216

[13] S.P. Chapoval, T. Neeno, C.J. Krco, E.V. Marietta, J. Harders, C.S. David, J. Immunol. 161 (1998) 2032

[14] M.P. Williamson, T.F. Havel, K. Wüthrich K., J. Mol. Biol. 182 (1985) 295

[15] K. Wüthrich, NMR of proteins and Nucleic Acids, John Wiley & Sons, New York (1986)

[16] J. Cavanagh, A.G. Palmer, W. Fairbrother, N. Skelton, Academic Press, San Diego (1996)

[17] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Proc. Natl. Acad. Sci. USA 94 (1997) 12366

[18] M. Salzmann, G. Wider, K. Pervushin, K. Wüthrich, J. Biomol. NMR 15 (1999) 181

[19] E. Becker, High Resolution NMR, Academic Press, San Diego (1999)

[20] D. Neuhaus, M. P. Williamson, The Nuclear Overhauser Effect, John Wiley & Sons, New York (2000)

[21] N. Tjandra, J.G. Omichinski, A.M. Gronenborn, G.M. Clore, A. Bax, Nat. Struct. Biol. 4 (1997) 732

[22] A.T. Brünger, X−Plor Version 3.1: A System for X−Ray Crystallography and Nmr, Yale Univ. Press, Yale (1993)

[23] S.K. Huang, P. Zwollo, D.G. Marsh, Eur. J. Immunol. 21 (1991) 1469

[24] S.K. Huang, D.G. Marsh, Immunology 73 (1991) 363

[25] M. Roebber, D.G. Klapper, L. Goodfriend, W.B. Bias, S.H. Hsu, D.G. Marsh, J Immunol. 134 (1985) 3062

[26] S. Dalal, S. Balasubramanian, L. Regan, Nat. Struct. Biol. 4 (1997) 548

[27] D.A. Vidusek, M.F. Roberts, L. Goodfriend, Biochemistry 24 (1985) 2747

[28]T. Rafnar, B. Ghosh, W.J. Metzler, S.K. Huang, M.P. Perry, L. Mueller and D.G. Marsh, J. Biol. Chem., 267 (1992) 21119.

[29]W.J. Metzler, K. Valentine, M. Roebber, M.S. Friedrichs, D.G. Marsh and L. Mueller, Biochemistry, 31 (1992) 5117.

[30]T. Rafnar, M.E. Brummet, D. Bassolino−Klimas, W.J. Metzler and D.G. Marsh, Mol Immunol., 35 (1998) 459.

[31]W.J. Metzler, K. Valentine, M. Roebber, D.G. Marsh and L. Mueller, Biochemistry, 31 (1992) 8697.

[32]1BBG; G.L. Warren, C.J. Turner, G.A. Petsko, A.T. Brunger

[33]G.A. Mueller, A.M. Smith, D.C. Williams Jr., G.A. Hakkaart, R.C. Aalberse, M.D. Chapman, G.S. Rule and D.C. Benjamin, J. Biol. Chem., 24 (1997) 26893

[34]G.A Hakkaart, R.C. Aalberse, R. van Ree, Allergy 53 (1998) 165

[35]S. Ichikawa, H. Hatanaka, T. Yuuki, N. Iwamoto, S. Kojima, C. Nishiyama, K. Ogura, Y. Okumura and F. Inagaki, J. Biol. Chem. 273 (1998) 356

[36]G.A. Mueller, D.C. Benjamin and G.S. Rule, Biochemistry 37 (1998)12707

[37]L. Holm and C. Sander (1996) Science 273 (5275) 595

[38]P. Bork, L. Holm, C. Sander, J. Mol. Biol. 242 (1994) 309

[39]P. Wiedemann, K. Giehl, S.C. Almo, A.A. Fedorov, M. Girvin, P. Steinberger, M. Rudiger, M. Ortner, M. Sippl, C. Dolecek, D. Kraft, B. Jockusch and R. Valenta, J. Biol. Chem. 271 (1996) 29915

[40]C. Dolecek, S. Vrtala, S. Laffer, P. Steinberger, D. Kraft, O. Scheiner, R. Valenta, FEBS Lett. 335 (1993) 299

[41]S. DeMarino, M.A. Castellione−Morelli, F. Fraternali, E. Tamborini, G. Musco, S. Vrtala, C. Dolecek, P. Arosio, R. Valenta, A. Pastore, Structure, 7 (1999) 943

[42]C. Faber, A. Lindemann, H. Sticht, A. Ejchart, A. Kungl, M. Susani, R.W. Frank RW, D. Kraft, M. Breitenbach and P. Rösch, J. Biol. Chem. 27 (1996) 19243

[43]A.J. Kungl, M. Susani, A. Lindemann, M. Machius, A.J. Visser, O. Scheiner, D. Kraft, M. Breitenbach and M. Auer, Biochem Biophys Res Commun. 223 (1996) 187

[44]M. Gajhede, P. Osmark, F.M. Poulsen, H. Ipsen, J.N. Larsen, R.J.J. van Neerven, C. Schou, H. Loewenstein and M.D. Spangfort, Nat. Struct. Biol. 3 (1996) 1040

[45]K. Schweimer, H. Sticht, J. Nerkamp, M. Boehm, M. Breitenbach, S. Vieths and P. Rösch, Appl. Magn. Reson., 17 (1999) 449

[46]P. Neudecker, K. Schweimer, J. Nerkamp, M. Boehm, H. Sticht, S. Scheurer, S. Vieths, & P. Rösch. J. Biomol. NMR (2000) in press

[47]P. Neudecker, Diploma Thesis, Bayreuth 2000

[48]D.R. Flower, A.C.T. North, T.K. Attwood, Prot. Sci. 2 (1993) 753

[49]A. Bufe, M.D. Spangfort, H. Kahlert, M. Schlaak, W.M. Becker, Planta 199 (1996) 413

[50]Y. Tsujishita, J.H. Hurley, Nat. Struct. Biol. 7 (2000) 408

[51]

[52]V. Tugarinov, A. Zvi, R. Levy, Y. Hayek, S. Matsushita, J. Anglister, Structure Fold. Des. 15 (2000) 385

[53]H. Takahashi, T. Nakanishi, K. Kami, Y. Arata, I. Shimada, Nat. Struct. Biol. 7 (2000) 220

[54] R.A. Sayle, E.J. Milner−White, Trends Biochem. Sci. 20 (1995) 374.

Table 1 – Sequence Alignment of Allergens

Ragweed allergens

```
Amb a 5  LVPCAWAGNV CGEKRAYCCS DPGRYCPWQV VCYESSEICS KKCGK
Amb t 5 DDGLC-YEGTN CGKVGKYCCS PIGKYC---- VCYDSKAICN KNCT
```

Dust mite allergens

```
Der f 2 DQVDVKDCAN NEIKKVMVDG CHGSDPCIIH RGKPFTLEAL FDANQNTKTA KIEIKASLDG
Der p 2 SQVDVKDCAN HEIKKVLVPG CHGSEPCIIH RGKPFQLEAV FEANQNTKTA KIEIKASIDG

Der f 2 LEIDVPGIDT NACHFVKCPL VKGQQYDIKY TWNVPKIAPK SENVVVTVKL IGDNGVLACA
Der p 2 LEVDVPGIDP NACHYMKCPL VKGQQYDIKY TWNVPKIAPK SENVVVTVKV MGDDGVLACA

Der f 2 IATHGKIRD
Der p 2 IATHAKIRD
```

Star-type allergens

```
Bet v 1 GVFNYETETT SVIPAARLFK AFILDGDNLF PKVAPQAISS VENIEGNGGP GTIKKISFPE
Pru a 1 GVFTYESEFT SEIPPPRLFK AFVLDADNLV PKIAPQAIKH SEILEGDGGP GTIKKITFGE

Bet v 1 GFPFKYVKDR VDEVDHTNFK YNYSVIEGGP IGDTLEKISN EIKIVATPDG GSILKISNKY
Pru a 1 GSQYGYVKHK IDSIDKENYS YSYTLIEGDA LGDTLEKISY ETKLVASPSG GSIIKSTSHY

Bet v 1 HTKGDHEVKA EQVKASKEMG ETLLRAVESY LLAHSDAYN
Pru a 1 HTKGNVEIKE EHVKAGKEKA SNLFKLIETY LKGHPDAYN
```

Table II – Summary of Solution Allergen Structures

Allergen	*Source*	*amino acids*	*Structure class*	*Fold type (54)*	*closest structural neighbors*	*Reference*	*pdb−entry*
Amb a 5	short ragweed *(Ambrosia artemisiifolia)* pollen	45	α−β	−	−	31	−
Amb t 5	giant ragweed *(Ambrosia trifida)* pollen	40	α−β	−	−	29 28 32	1BBG
Der p 2 (mutant)	birch pollen *(Betula pendula)*		all−β	immunoglobin	rho, *homo sapiens* (1RHO, GDP dissociation inhibitor)	33 36	1A9V
Der f 2 (mutant)	dust mite *(Dermatophagoides Pteronyssinus)*	129	all−β	immunoglobin	rho, *homo sapiens* (GDP dissociation inhibitor)	35	1AHK
Phl p 2	dust mite *(Dermatophagoides Pteronyssinus)*	129	all−β	Calcium/lipid binding domain	fibronectin folds	41	1BMW
Bet v 1	birch *(Betula pendula)* pollen	159	α−β	Tata−box binding protein like (β−α−β(4)−α)	mln64, *homo sapiens* (1EM2, cholesterol transport)	42 44 45	1B6F 1BTV
Pru a 1 (wt)	cherry *(Prunus avium)*	159	α−β	Tata−box binding protein like	mln64, *homo sapiens* (1EM2, cholesterol transport)	46 47	1E09

Figure captions

Fig. 1: The Rasmol [54] picture of the Amb t 5 (pdb−entry: 1BBG) structure shows the clearly defined helix and the short two−stranded antiparallel β−sheet. Indicate is loop 3 that is proposed to contain the epitope.

Fig. 2: The structure of Der p 2 (pdb−entry: 1A9V) shows clearly the similarity of the fold with the immunoglobulin type of folds (Rasmol [54] picture).

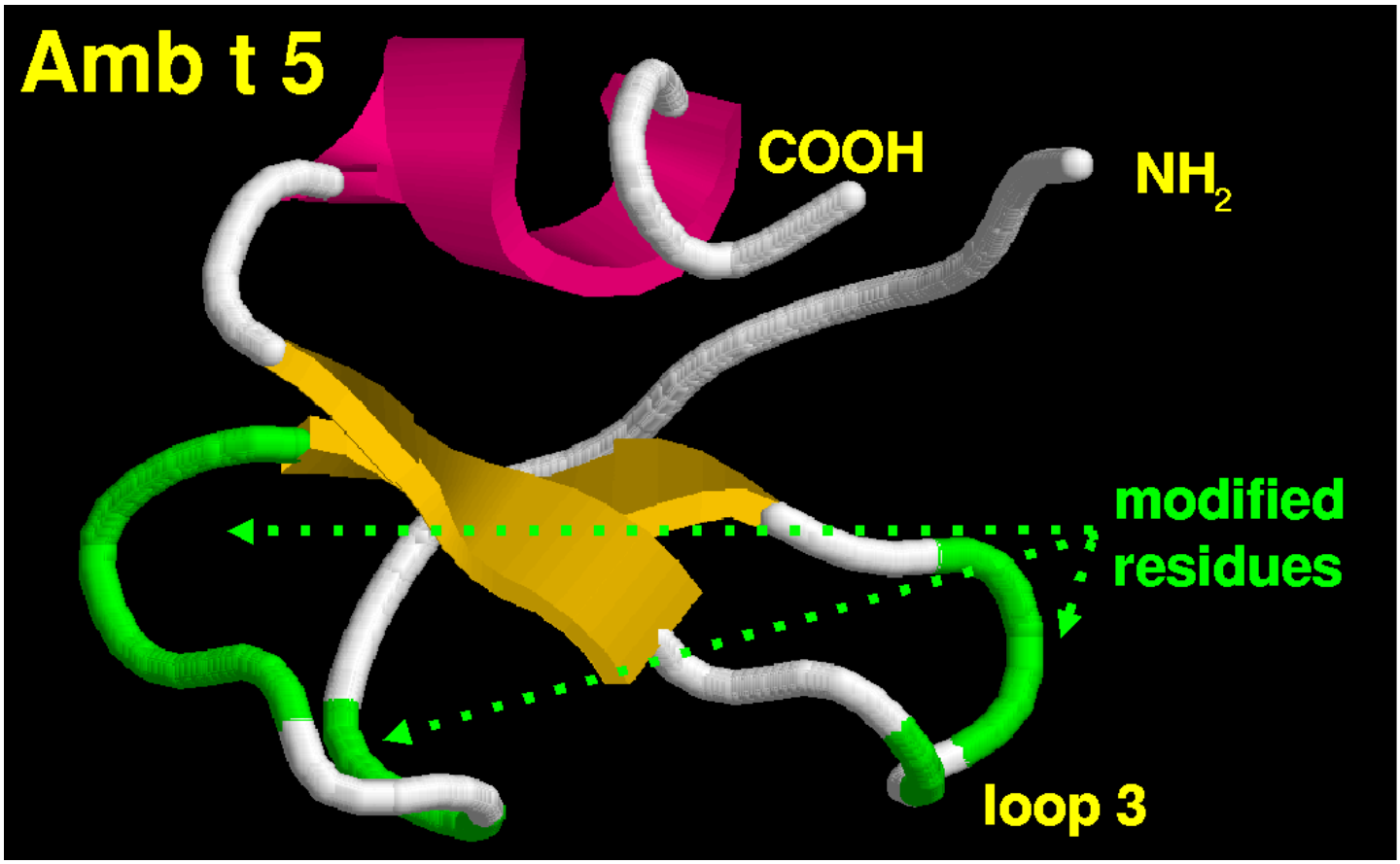
Fig. 3: The structure of Phl p 2 (pdb: 1A9V). The opposed β−sheets similar to the immunoglobulin fold can easily be seen. (Rasmol [54] picture; pdb−entry: 1WHO was used for the figure as the secondary structure elements are better defined according to Rasmol in this X−ray structure)

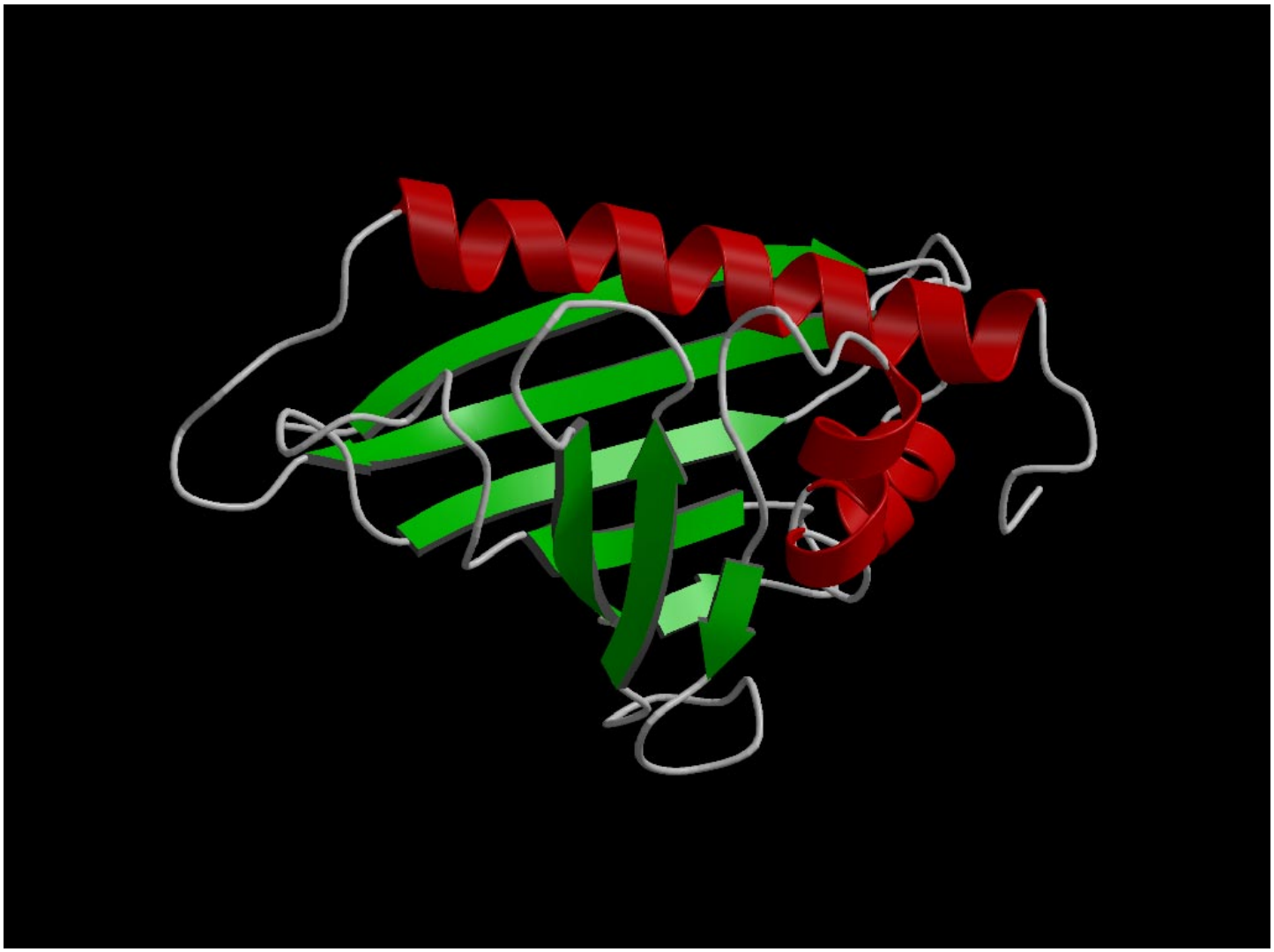
Fig. 4: The Rasmol [54] picture of the Pru av 1 structure (pdb−entry: 1E09) shows the 7 β−strands in a basket−type arrangement. This basket is closed by α−helices 1 and 2 in combination with the long COOH−terminal α−helix 3.

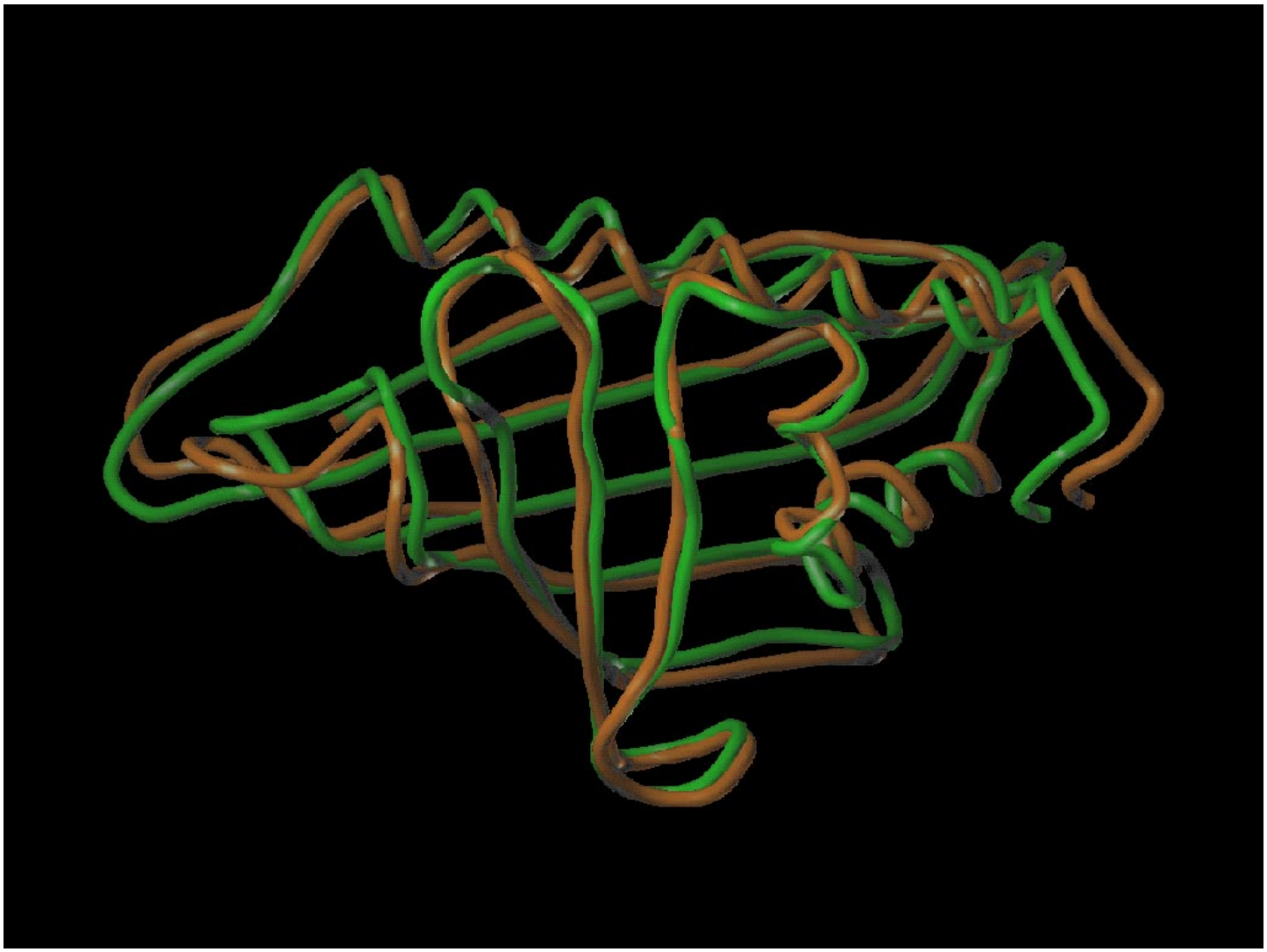
Fig. 5: Backbone overlay of Pru av 1 (pdb−entry: 1E09) and Bet v 1 (pdb−entry:1BV1, X−ray structure) showing the near−identity of the two protein conformations (SYBYL, Tripos Inc.)

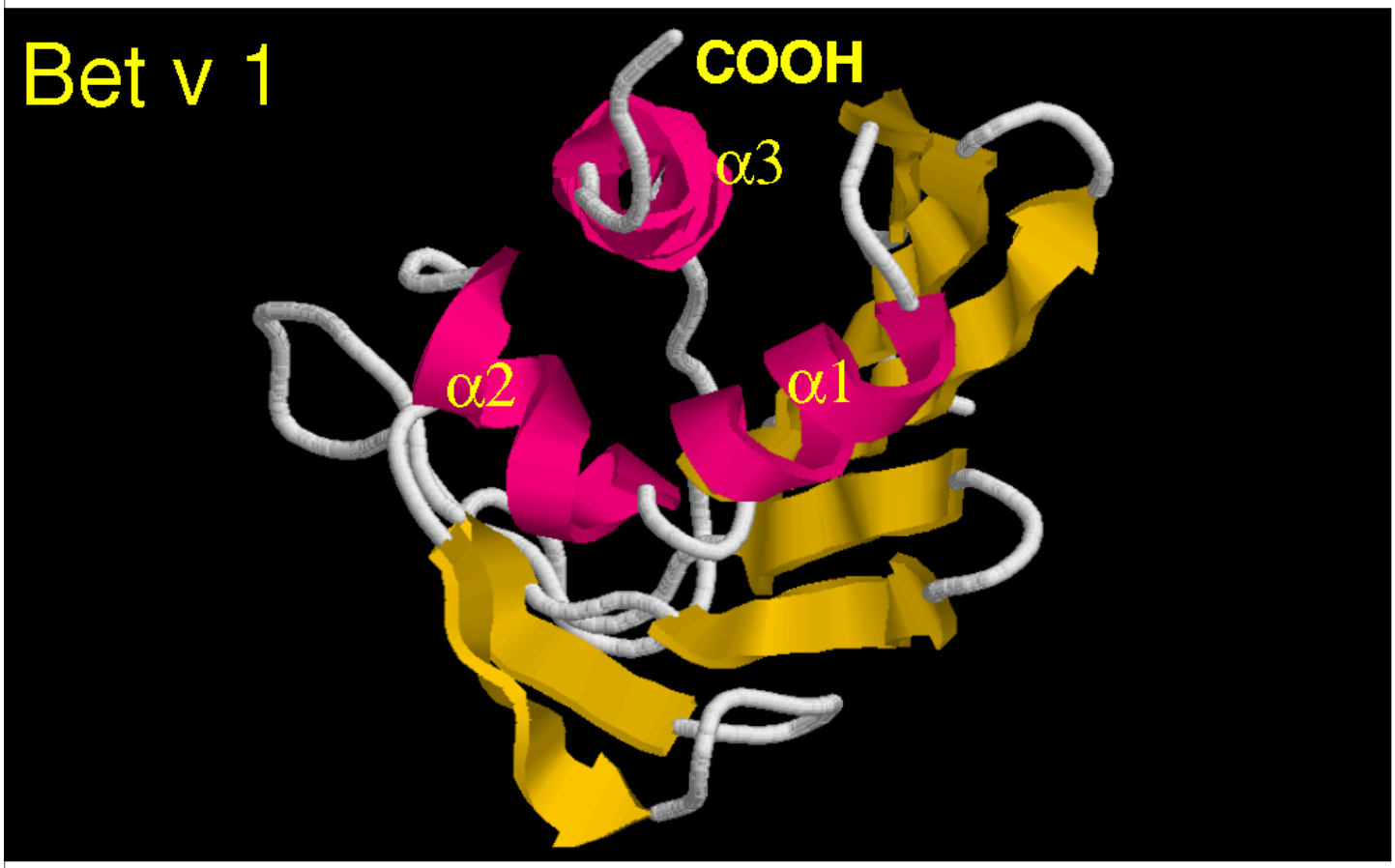
Fig. 6: The Rasmol [54] picture of Bet v 1 (pdb−entry: 1B6F) oriented along the COOH−terminal helix shows the basket−type arrangement of β−strands and the lid−clip arrangement of the α−helices that is also found in the structure of the human cholesterol binding protein mln64 [50]

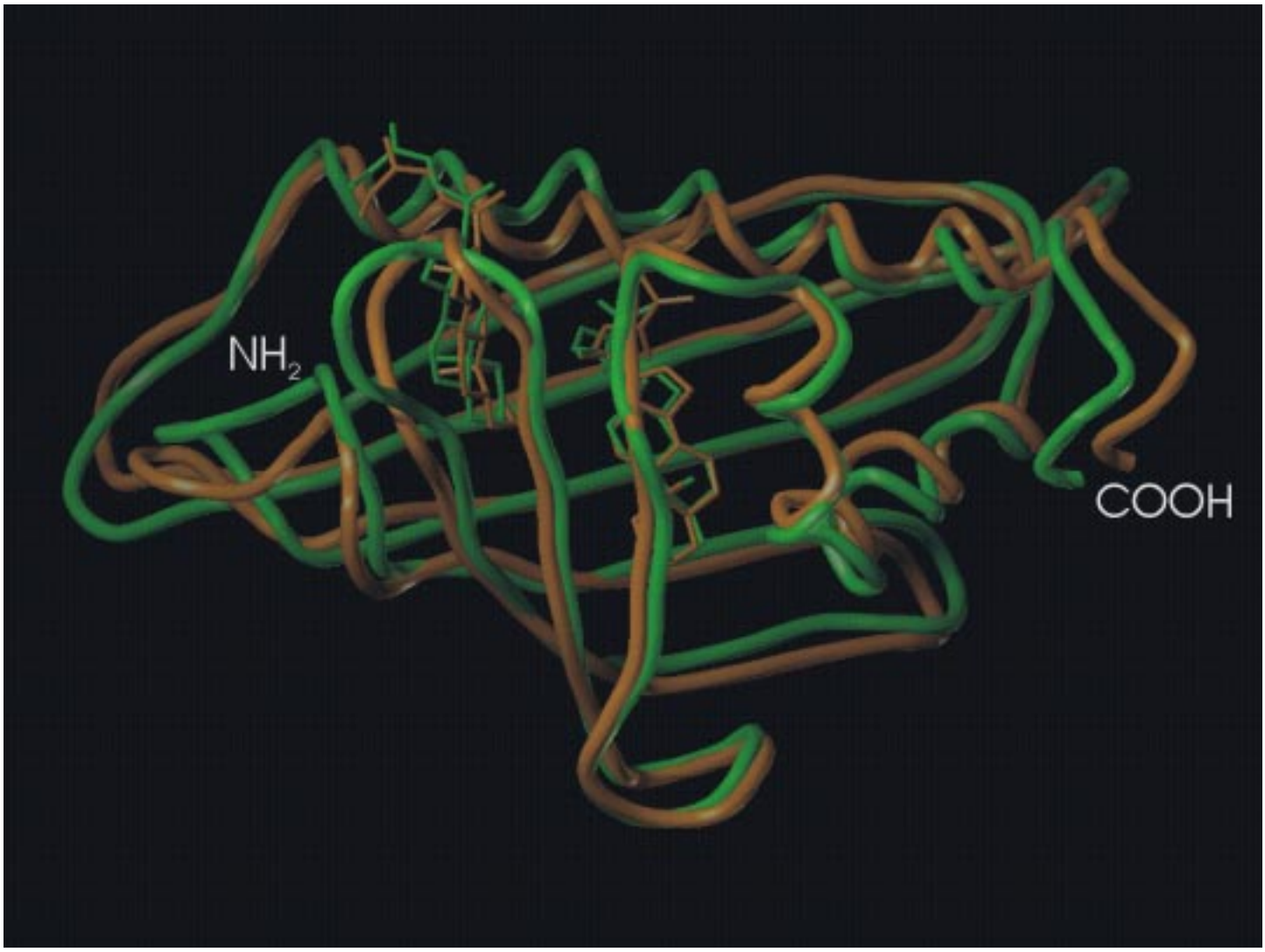
Fig. 7: A combination of molecular modeling and NMR−experiments suggests that two plant steroid molecules may bind in the hydrophobic cavity of Pru av 1 [51].


Amb t 5
COOH
NH2
modified
residues
loop 3






Bet v 1
COOH
α3
α2
α1


NH2
COOH