Chapter 3

Intermediate States of Fetuin and Asialofetuin at Acidic pH

Work from this chapter published in:

1.3.4 Hydrophobic collapse

The hydrophobic collapse model, [44] the hydrophobic effect is considered to be the driving force for folding, squeezing out water in a nonspecific manner, and the subsequent rearrangement of the collapsed state is the rate limiting step. This model predicts an intermediate state that has been called molten globule, which has been characterized both kinetically and at equilibrium as an expanded form of the native state [45,46]. The protein α-lactalbumin forms a molten globule state at low pH known as the A-form. NMR has shown that the A-form contains at least 2 of the helices present in the native form and similar intermediates form during folding.

1.3.5 Nucleation condensation mechanism

Folding times [4] which are in good agreement with those observed in experiments on small proteins were provided by the nucleation-condensation mechanism [5, 47-49], which largely unifies the classical nucleation-growth mechanism with the hydrophobic collapse model. In contrast to the well-defined protein-folding nucleus in the nucleation-growth mechanism [43], the folding nucleus of the nucleation-condensation mechanism is diffuse and initially consists of several neighboring amino acid residues whose conformations are stabilized by long-range interactions with residues that are remote in sequence. An essential component of this mechanism is the concurrent formation of secondary and tertiary structures [5, 47, 49, 50,51].

1.3.6 The Framework model

Framework model, [52] envisages the formation of secondary structure elements followed by the docking of those elements to form tertiary interactions in the rate limiting step. The framework hypothesis was boosted when it was found that protein fragments corresponding to secondary structure elements could be partly folded in the absence of tertiary interactions [53,54]. Folding begins by formation of 'flickering clusters' of secondary structure elements which are stabilized by packing together leading to complex tertiary structures [55-57]. One problem with this theory is that peptides do not generally form stable secondary structures in solution. The highest equilibrium constant for a coil-to-helix transition is 1.3 for poly-L-alanine. Studies on poly-L-glutamate showed that the
time scale for α-helix formation is ~ 10⁻⁷ seconds i.e. rapid on the time scale of folding. Early experiments in which fragments of proteins were tested for α-helix formation proved negative. However, it was recently found that a peptide fragment of ribonuclease A (C-peptide; 13 residues) forms a stable α-helix. This is probably due to specific side chain interactions seen in the crystal structure of ribonuclease A. NMR has been used to demonstrate framework intermediates early in the folding of several proteins. This technique involves refolding the protein in D₂O using rapid-reaction techniques and relies on the fact that H and D nuclei can be distinguished by NMR. Initially the protein is denatured in D₂O which will cause all of the main chain >N-H groups to exchange their H atoms with D₂O, i.e. they become >N-D. A sample is rapidly transferred to a refolding buffer and a pulse of H₂O is added after a certain time interval. H-D exchange is then quenched by transfer to low pH, folding is allowed to complete and the final NMR spectrum is recorded. Repeating the experiment with the H₂O pulse at different time intervals after the initiation of folding makes it possible to detect which residues are buried at different stages of folding. Residues which are exposed at the time of the H₂O pulse will exchange their main chain >N-D groups for hydrogen (from the H₂O) and this can be detected in the final NMR spectrum. In contrast, those residues buried in the folding intermediate cannot exchange rapidly and their hydrogen NMR signals will be absent from the final spectrum. This technique has shown that residues involved in the β-sheet of ribonuclease form hydrogen bonds early in folding (1.5 seconds out of 10 min refolding time). Similarly, residues of the amino and carboxy terminal helices of cytochrome c fold early. These helices interact via a hydrophobic patch indicating that helix contact may occur as an early event.
1.3.7 Energy landscape and funnel theory

The term "new view" stresses the possibility of a diverse "myriad of pathways" with "delocalized" transition states. The central feature of the new view is the replacement of the pathway concept with picturesque funnel diagrams to illustrate features of protein folding and the role of ensembles [58, 59]. In this view, energy landscapes are used to describe the kinetics and thermodynamics of the folding reaction. These landscapes are funnel-shaped, with slopes displaying varying degrees of roughness [58, 60].

The drive for a protein to fold to its native state originates from a strong slope of the energy landscape toward native conformations. However, the roughness of the energy landscape could render the folding reaction less effective, a phenomenon called frustration, the concept of frustration reflects the inability of a protein to energetically satisfy all its interactions in any given conformation. Frustration in protein folding landscape arises both from topological and sequence specific energetical traps inclining proteins to accumulate intermediates during folding, rendering the process less efficient [61]. The most stable state is then the structure maximizing the fraction of attractive interactions, which is the minimally frustrated structure. Given a set of coordinates describing the dependence on energy of a conformation, the free energy of the solvated protein as function of these coordinates defines a hypersurface, or energy landscape. A
conformational transition of the protein, like the folding event, is then given by a specific trajectory or bundle of trajectories on this energy landscape. A smooth energy landscape allows rapid transitions, whereas a rugged energy surface has kinetic traps, which slow down folding. Frustration is associated with a rugged energy landscape.

The picture offered by the Levinthal paradox is, in terms of landscape, a flat energy surface with a single narrow absolute minimum representing the native state, similarly to a golf course with a single hole. A random search through the flat surface, without any bias, would be successful in finite time only with an extremely low probability. A different extreme is represented by a rugged landscape with many maxima and minima without a preferred conformation: this would be the case of a random heteropolymer. The energy landscape of a folding protein is described as a rugged surface sloped down from all directions into a global minimum, representing the native state. It is explicitly assumed that there is a kinetic flow through a series of states that are progressively lower in energy. At the top of the funnel the protein exists in a number of random states that have relatively high entropy and high enthalpy. Progress down the funnel is given by collapse and reconfiguration. The reconfiguration occurs as a motion through adjacent, thus geometrically similar, conformations, by a Brownian-like motion. The acquisition of native structure reduces the free energy and drives the protein towards the bottom of the funnel, while enthalpy is progressively reduced. The driving force is working against entropy, which decreases simultaneously with entropy.

The progress of folding can be described by the parameter $Q_i$, the fraction of native contacts in the state $i$:

$$Q_i = \frac{C(i)}{C(n)}$$

where $C(n)$ is the number of contacts formed by residue pairs in the native state and $C(i)$ is the number of such contacts conserved in state $i$. The molten globule state is rather compact but still highly non-native. The transition state is located at a higher value of $Q_i$, thus closer to the native structure. If $Q$ does not increase uniformly, but in few large steps, then the protein folding occurs through nucleation. After passing through the transition state, there is a single low energy state corresponding to the native conformation. This drives the protein to fold completely, while the absence of other relevant minima prevents from misfolding.
The folding funnel. The height of the funnel represents enthalpy, which decreases going downhill (here referred to as energy, left arrow). The width of the funnel represents entropy. On the right the decrease of the fraction of native contacts $Q$ is given by an arrow. The molten globule represents a compact denatured state. It is separated from the native state by an energy gap which guarantees the stability of the native structure.

1.4 Implications of protein folding studies

- Although techniques like X-ray crystallography and NMR are used for determination of three-dimensional structures of biological macromolecules but these studies are not enough to predict the structure of proteins. Folding studies are a prerequisite to understand the structure-function relationships of proteins. The challenge of the protein-folding problem is to have the ability to predict protein structure more accurately. To be able to predict the structure of the protein from just the amino-acid sequence would have tremendous impact in all of biotechnology and drug design.

- Solving the folding problem has enormous implications exact drugs can be designed theoretically on a computer without a great deal of experimentation. For the pharmaceutical industry, this holds the prospects of greatly reducing the cost and expense of developing new therapeutic drugs.
Genetic engineering experiments to improve the function of particular proteins will be possible. One of the major areas of study these days is designing proteins with novel biological function and for this purpose one should know how each amino acid contributes to the structure and native fold. Modification of proteins can also be carried out by altering the primary sequence and seeing the consequent effect on the structure and folding. Expression of recombinant genes leads to the production of inclusion bodies that are insoluble and inactive forms of expressed protein. These proteins have to be made active for which protein folding studies are required.

The process of protein folding is remarkably efficient, but sometimes it can go wrong. This can have harmful consequences, as the incorrect folding of proteins is thought to be the cause of diseases, such as Alzheimer’s disease and cancer. Protein folding studies help in better understanding of such diseases caused due to aggregation and misfolding of the native functional protein into an inactive disease causing form.

Formation of amyloid fibrils causes disorders such as scurvy, cystic fibrosis etc. The amyloid and prion diseases appear to result from conversion of one of the soluble and functional proteins into β-sheet rich quaternary structures that are often fibrillar and this occurs at low pH or under denaturing conditions. Protein folding studies also help to understand how some human proteins undergo conformational changes that render them pathogenic.

1.5 Protein folding diseases

Protein folding diseases can be divided into two groups: in the first, excessive quantities of wrongly folded proteins collect in the form of uncontrolled piles of molecular rubbish. This is the group of diseases known as amyloidoses, of which Alzheimer’s disease is the best-known example. In the other, a small error in the genetic blueprint leads to incomplete folding of a protein, which affects its function. This might, for instance, happen to p53 the malfunctioning of this central tumor suppressor could cause cancer.
1.5.1 Amyloidoses

The common characteristic of all amyloidoses is the collection of plaques of insoluble protein in the extracellular tissue, which cannot be broken down by enzymes. Their ordered structure gives them crystal-like properties: they are made up of long filaments (fibrils) that are formed from densely packed β-pleated sheets of identical proteins. There are about 20 different proteins that can act as the building blocks of these fibrils, each of which is associated with a different disease.

In so-called systemic amyloidoses, the precursors of these plaques are transported through the bloodstream from their point of origin to their point of deposition. Localized amyloidoses are of greater clinical significance, as they mainly affect the central nervous system, the extracellular tissue of which is particularly susceptible to damage.

1.5.2 Alzheimer’s disease

One of the main characteristics of Alzheimer’s disease is the accumulation of plaques of insoluble β-amyloid in the brain. It is still not certain whether these plaques are a cause or a consequence of the disease, but there is a lot of evidence for the former being the case. The β-amyloid plaques are formed by cleavage of amyloid-precursor protein (APP) by two different enzymatic activities, which release amyloid-b peptide fragments that are 40 or 42 amino acids long. These then form fibrils, which aggregate into insoluble clumps of β-amyloid plaques that surround neurons and might cause damage. But this cleavage also occurs in healthy individuals and soluble β-amyloid proteins are normal constituents of brain tissue.

How, then, do the plaques form in Alzheimer’s patients? It is thought that the misfolding of the protein dramatically alters its properties. In the normal protein, hydrophobic (water-repelling) amino acids bury themselves inside the protein right from the start of folding. However, if the protein folds wrongly, these hydrophobic amino acids are exposed and they rapidly seek out and bind to hydrophobic groups on other protein molecules, forming the insoluble aggregates or plaques that are found in Alzheimer’s patients.
1.5.3 Prion diseases

Transmissible spongiform encephalopathies (TSEs), which include mad cow disease (bovine spongiform encephalopathy; BSE) and Creutzfeld Jakob disease (CJD) in humans, are special forms of amyloidosis in which the victim's brain degenerates to a structure that looks like a porous sponge. These conditions seem to occur when normal human protein particles called prions misfold. The normal human prion is a component of the membrane of healthy nerve cells (called PrP\textsuperscript{C}), which folds properly, remains soluble and is disposed of without problem. It can, however, misfold in a particular way, which allows it to take on an infectious, incorrectly folded three-dimensional form (called PrP\textsuperscript{SC}), presumably due to a genetic mutation. The infectious prion, which can be transmitted in the diet, triggers a domino effect in healthy prions, forcing them to adopt its incorrectly folded form.

1.5.4 Misfolding and Cancer

Whereas too much of an incorrectly folded protein can cause amyloidoses, another group of protein folding diseases is caused by lack of a correctly folded protein. This form of protein folding defect is thought to be involved in diseases such as cystic fibrosis, but mainly affects a protein called p53, which occupies the most important position in the body’s cancer resistance network. Normally, the p53 system is switched off or, at most, is in stand-by mode. It is activated inside a cell if the cell becomes excessively stressed or damaged, which can lead to genetic mutations in DNA that can cause the uncontrolled division and proliferation of cells that is the hallmark of tumor formation. p53 is so good at its job that even a single break in the DNA strand is enough to activate it. It rushes into the cell nucleus and induces the production of other proteins that stop uncontrolled cell division or trigger the programmed death of a cell. This tumor-suppressing function of p53 is so important that the protein has been described as the guardian of the genome. So, it is no surprise that faults in the p53 gene can be disastrous. Even a mutation in one of the letters (nucleotides) in the gene can be enough to lead to the expression of p53 proteins that do not fold correctly. Half crippled, they cannot carry out their job properly, so the damage to the DNA that would normally be repaired goes unnoticed, allowing the
abnormal cell to grow in an uncontrolled manner. This type of mutation in p53 is thought to occur in 50% of all cases of cancer and as many as 95% of all cases of lung cancer.

1.6 Intermediates in protein folding pathways

The native state of a globular protein is defined as the small ensemble of compact conformations, reached under folding conditions in vivo or in vitro, in which the protein is stable and performs its action. This state is characterized by a low amount of entropy, because all protein atoms are kept by mutual interaction in a well-defined geometry. Intramolecular enthalpy is also relatively low because many attractive interactions are satisfied. Solvent molecules surround the folded protein in its native state. Native conformation can be disturbed by perturbing the folding conditions in some way, like increasing temperature, changing the solvent pH or adding a chemical denaturant. The most commonly used chemical denaturants are urea and guanidinium chloride.

Both denaturants favor the denatured state by increasing the solubility of the unfolded chain in an aqueous solution. In 8M urea and in 6M GndHCl proteins are thought to approach a random coil state. The exact mechanism of how this is accomplished is still under debate. One of the theories is the denaturant binding model. In this model the denaturant molecules bind to peptide groups and side chains and thereby create a “shell” around the folded and the unfolded polypeptide chain. As the concentration of denaturant increases the unfolded state will be favored since it has a larger surface area exposed to the solvent, more molecules can bind. The driving force for denaturation with chemical denaturants is therefore proportional to the change in exposed surface area going from the native to the denatured state.
Out \[ N \]

In contrast to temperature denaturation, chemical denaturation is often a reversible process. This is possible since the hydrophobic groups of the unfolded chain are shielded by the denaturants, which prevent aggregation. The fact that the process is reversible and that the protein approaches a maximum unfolded state is important when we want to measure conformational transitions.

It is difficult to characterize the structural properties of the denatured state, since it strongly depends on the unfolding conditions. One can describe the ensemble by means of parameters like the radius of gyration or the hydrodynamic volume, which both define the average compactness. Experimental techniques like infrared and circular dichroism spectroscopy provide information on secondary structure elements. When a high percentage of secondary structure is lost, which is the case under strong unfolding conditions, the unfolded state assumes a random coil structure.

The denatured state is characterized by high conformational entropy, because native interactions are lost and residues are free to assume a large set of arrangements. Intramolecular enthalpy may be higher because of loss of native interactions. The solvent molecules get in contact with hydrophobic parts of the protein chain, which were buried in the native state. This fact causes water molecules to assume more ordered conformations in order to minimize the contact with nonpolar groups, what reduces the entropy of the solvent relatively to the folded state. Therefore, one can understand the entropy increase as a driving force for folding. When restoring the folding conditions in vitro, the spontaneous refolding of a denatured protein is usually initiated. This means that the molecule is able to find the native state spontaneously. As proteins function they fold-unfold in specific ways. How proteins can quickly refold without tangling remains a central mystery in molecular biology.

To understand fully the conformational behavior of a protein, it is necessary to define not only the structure of its native state but also that of various intermediate states.
Knowledge of the latter is essential because they are the starting points of the process by which folding to the active form of the protein is achieved following biosynthesis. Study of the intermediate states of the protein folding has entered a new stage, aided by recent technical advances. There has been increasing emphasis on the biological importance of non-native conformations of protein ranging from various denatured conformations to aggregated forms [62-68]. Proteins may unfold through intermediate(s), which retain(s), a substantially high content of secondary structure but little or no tertiary structure known as molten globules (MG) [24,69].

The MG state attracts special attention; it may have some features of the native fold. However, this state differs from native state by the absence of close packing throughout the molecule and by a substantial increase in fluctuations in the side chains as well as of larger parts of the molecules [70-72]. Recently, increasing evidences support the idea that MG state besides having secondary structure may posses well defined tertiary contacts as well [73-76]. Thus, the structural similarity between the MG state and native protein may have a significant bearing in understanding the protein folding problem [77]. Studies have revealed the biological significance of MG states in processes such as chaperon binding [78,79] and transport across membranes [62,80]. A full understanding of the mechanism of protein folding requires the knowledge of the structures, relative energetics, and dynamics of the species populating the folding pathway.
PROTEIN FOLDING
KINETIC AND THERMODYNAMIC MEASUREMENTS

Y is observable:
- $A_{280}$
- Fluorescence Int.
- viscosity
- CD
- etc.

KINETICS
Timed measurements

THERMODYNAMICS
Equilibrium measurements

TRAPPED INTERMEDIATES

- disulfide
  - trap incorrect disulfides by labeling folding protein with iodoacetamide

- cis/trans X-Pro bonds
  - some proteins with cis X-Pro bonds, isomerization on N $\rightarrow$ D; must revert to cis on D $\rightarrow$ N

- molten globule
  - some protein at low pH
  - some Ca binding proteins w/o Ca

- secondary struct
  - deuterate protein, initiate folding, then pulse in H2O. Amide D's in secondary struct. protected. Do NMR
1.7 Lectins

1.7.1 What are lectins?
Lectins are proteins of non-immune origin that bind carbohydrates in a reversible manner, without having any enzymatic activity towards this bound carbohydrate. The term lectin was first coined by Boyd & Shapleigh, 1954, [81], and is derived from the latin word *legere*, which means "to choose". Lectins were originally discovered through their ability to agglutinate blood cells via cross linking of the sugar residues present on their surface. Lectins are ubiquitous in nature and can be found in viruses, bacteria, plants and animals, including humans. They are the natural partners of the various oligosaccharides that nature uses to "tag" cells via glycoproteins and glycolipids. Lectins play a role in important processes like infection, host defense, fertilization, cancer, protein transport and embryogenesis [82]. They are also important as tools in biology, histology and medicine, e.g., as an aid in the purification of glycoproteins or the determination of blood groups [83]. Lectin-carbohydrate recognition is one of the three fundamental biological recognition mechanisms, the two others being protein-nucleic acid recognition and protein-protein recognition. A detailed understanding of lectin-carbohydrate recognition mechanisms is therefore of great practical as well as fundamental value.

1.7.2 The legume lectins
The legume lectins are a family of sugar binding proteins found in the seeds, and in smaller amounts in roots, stems, leaves and bark of plants belonging to the *Leguminosae* family [84, 85]. The exact function of the legume lectins *in vivo* has not yet been established although considerable attention has been devoted to the possibilities that they may be involved in the defense of plants against predators or in the interaction of the plant with *Rhizobium* symbionts [86-88]. Other functions of lectins in plants may include: enzymes (but unknown substrate), storage of proteins, cell wall extension, mitogenic stimulation, transport of carbohydrates, packaging and/or mobilization of storage materials. Biological activity of the lectins may be attributed to the metal ions which are the essential part of the native structure of most leguminous lectins. The stability of the native structure of most lectins is thought to be caused by the hydrophobic interactions.
Recently, homologues in other plant families [89, 90] and in animals have been described [91, 92]. They have been used for decades as a model system for the study of protein-carbohydrate interactions, because they show an amazing variety of binding specificities and are easy to obtain and purify. Over the years, a quite impressive amount of structural data has been gathered. For reviews with an emphasis on structural features, see [85, 93-95].

A typical legume lectin monomer, lentil lectin, PDB entry 1LES [96]. β-Strands are shown as arrows. The conserve monosaccharide binding site is located on the top of the monomer, in a shallow depression. The bound Glc molecule is shown as a ball-and-stick model. The four sugar binding loops are shown in different colors (Ala-Asp cis-peptide bond in blue; metal binding loop in yellow; Ω-loop in orange; monosaccharide specificity loop in red). The metals are shown as gray spheres (Ca2+ dark gray; Mn2+ light gray).

Legume lectins are dimeric or tetrameric hetero- or homo-oligomers. The structures of the different legume lectin monomers are extremely similar, and their structure can be described as a β-sandwich consisting of a flat, six stranded sheet (the back sheet), and a curved, seven stranded sheet (the front sheet). A small third β-sheet, consisting of five strands, holds the two larger sheets together [97]. The sugar binding site is located on the concave side of the β-sandwich, formed by the curved front sheet, next to a double metal binding site. The topology of the monomer is related to the right-handed class 1 jellyroll fold [98] but it contains three insertions, one of which is directly involved in sugar and metal binding. Essentially the same topology is found in the galectins [99, 100] the pentraxins [101] and the spermadhesins [102, 103] all of which are lectins. Many other
Sugar binding or processing proteins have a topology related to the jellyroll fold [104], these include e.g. β-D-glucanase [105] PNGase F [106, 107] and cellulbiohydrolase I [108]. Sugar binding always seems to occur on the concave side of the β-sandwich, on the side of the front sheet in the legume lectins [109, 107]. The combination of this curved β-sheet and the connecting loops that fold over the sheet is especially suited for carbohydrate binding. The similarity between the topology of the legume lectins, the pentraxins, the galectins and the spermadhesins which have no significant sequential identity may thus be the result of convergent evolution. Moreover, the β-sandwich architecture may also be especially suited for generating a variety of quaternary structures, which is important for binding of multivalent epitopes and cross-linking of ligands.

Topology diagram of a typical legume lectin monomer (lentil lectin, PDB entry 1LES). β-Strands are shown as arrows, connecting loops as simple lines. The monomer consists of a flat six strand β-sheet (shown right), a curved seven strand β-sheet (shown left) and a small connecting five strand p-sheet (shown in gray). The topology is related to the jellyroll fold, with three additional inserts (shown as hatched strands). One of these inserts contains two of the four conserved sugar binding residues (Phe 123β, Asn 125β) and the metal binding loop (shown in bold). The two other conserved residues are Asp 81β (involved in the Ala-Asp cis-peptide bond) and Gly 99β (present in a large β-loop that packs against the front sheet). Ala 30α and Glu 31α belong to the monosaccharide specificity loop. Lentil lectin belongs to the Glc specific **Vicieae** tribe lectins and is proteolitically cleaved at position 181 (shown as a dotted line), resulting in a two-chain lectin (chains α and β).
1.7.2.1 Quaternary Structure

One of the unique features of the legume lectins is their variable quaternary structure, although the structures of their monomers are highly similar, they can associate in a number of different tetramers and dimers.

The so-called canonical dimer consists of two monomers that associate via the formation of a continuous twelve-strand β-sheet out of two flat back sheets, along the length of the dimer. This quaternary structure is adopted by pea lectin [110], favin [111], lentil lectin [112], the *Lathyrus ochrys* isolectins [113] and the α-amylase inhibitor from *Phaseolus vulgaris* [114].

Con A, the first legume lectin structure reported [115] consists of two canonical dimers that pack against each other with the central parts of their continuous twelve strand β-sheets. The β-strands of the two dimers are at an angle of almost 90°, and the interface consists mainly of electrostatic interactions. The Con A tetramer has been crystallized in different space groups. These structures show that the Con A tetramer is quite flexible, small rotations and/or translation of the subunits are possible [116, 117]. This is also illustrated by a naturally occurring "mutant" of Con A from *Canavalia brasiliensis*, whose quaternary structure differs slightly from the other Con A structures, presumably due to a mutation (Asp58Gly) at the dimer-dimer interface [118].

Peanut agglutinin [119, 97] forms a tetramer that can be considered as an association of two GS4 dimers. These two GS4 dimers associate via the formation of a canonical dimer, formed by one monomer from each GS4 dimer. The peanut tetramer is unique because it is the only known homo-tetramer that does not possess 222 nor 4 fold symmetry.
Different quaternary structures of the legume lectins. The known legume lectin tetramers can be considered as dimers of dimers, so the tetramers and their corresponding dimers are vertically arranged for ease of comparison. Con A contains two canonical dimers. Peanut agglutinin contains one canonical dimer and two GS4 dimers. There is no tetramer known that contains the EcorL lectin type dimer.

Remarkably, lectins that have no significant sequence identity with the legume lectins can form similar multimers. The galectins are Gal-specific mammalian lectins that possess the same fold as the legume lectins, although the position of the sugar binding site differs from the legume lectins. The dimers of a human galectin [99] and bovine spleen galectin [100] resemble a canonical legume lectin dimer, although two continuous \( \beta \) -sheets are running along the length of the dimer instead of one. Two lectins with an identical specificity, but with a different quaternary structure will necessarily form different cross-linked lattices with the same epitope.

Hence, quaternary structure can be used as a tool to obtain the correct cross-linked lattice or to maximize the avidity towards structures bearing multiple epitopes. Owing to their ability to differentiate between complex carbohydrates based on subtle variations, plant lectins are extensively used as probes for cell surface receptor sites. It is for this reason that the plant lectins (though first used as tools in the beginning of the century) continue to play a central role in immunology, cancer biology etc. and their importance has not
been affected by the advent of newer technologies. Specificity of peanut agglutinin has been modified so that it recognizes the T-tumor associated Thomsen-Friedenreich (Galβ 3GalNAc, T-antigen) alone at the expense of the structurally and conformationally related sugars viz. lactose (Galβ 3Glc) and N-acetyllactosamine. Discovery of a monomeric molten globule state for PNA with retention of biologic activity indicates that the monomers of legume lectins have the ability to bind sugars and that their oligomerization imparts them enhanced stability and topology necessary for their biologic activities.

1.7.3 Lectins and Cancer

Lectins present on the surface of tumor cells are targeted for therapeutic purposes. It has been found that treatment with anti-lectin antibodies can suppress growth of tumor cells in agarose, and inhibit lung colonization in vivo. Lectins have the potential use in cancer treatment strategies due to the fact that lectins present on the surface of tumor cells are capable of binding exogenous carbohydrate-containing molecules and internalize them by endocytosis. For example, wheat germ lectin (WGA) is found to induce lectin-dependent macrophage-mediated cytotoxicity against human bladder cancer (T-24) cells. Alveolar macrophage (AM) are phagocytes, mainly present in the pulmonary alveoli, are important in the antitumor defense mechanism of the lung because they can bind to the target cell-but are unable to induce cytolysis. However, studies have revealed that human AM tumorcidal activity can be induced by wheat germ lectins. Another finding is that the sensitivities of six human tumor cell lines depend on the number of receptor sites exist on the surface of WGA. Although the effector mechanism is still unknown, the binding of AM with tumor cells initiated by WGA may increase sensitivity to the cytotoxicity mediated by human AM. In addition, WGA is found to enhance the cell killing ability of murine peritoneal macrophages. In vivo studies show that WGA has an inhibitory effect on the growth of murine tumors. The tumorcidal activity of human blood monocytes can be induced by the WGA. As a result, the monocytes are able to become cytotoxic to four different human tumor cell lines: T-24 bladder carcinoma, A-375 melanoma, ACHN renal carcinoma, and U373MG glioblastoma.