Acta Biochimica Polonica

Vol. 46 No. 3/1999

487–508

QUARTERLY

Review

Folding initiation sites and protein folding*

Michał Dadlez

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5A, 02-106 Warszawa, Poland

Received: 23 June, 1999

Key words: protein folding initiation, residual structure, β-hairpin, α-helix

The paper discusses the role of local structural preferences of protein segments in the folding of proteins. First a short overview of the local, secondary structures detected in peptides, protein fragments, denatured proteins and early folding intermediates is given. Next the discussion of their role in protein folding is presented based on recent literature and data obtained in our laboratory. In conclusion it is pointed out that, during folding, local structures populated at low levels in denatured state may facilitate the crossing of the folding transition state barrier, and consequently accelerate the rate limiting step in folding. However, the data show that this effect does not follow simple rules.

Protein folding completes the fundamental process of the flow of information from gene sequence to protein structure and function. Thanks to enormous effort of many research groups during the last 30 years the main outline of the process begins to become clear. This review summarises recent developments in the *in vitro* experimental studies on the initiation of protein folding and the impact of early-forming structures on the efficiency of the entire process. Folding of proteins *in vivo* may require the support of a translation system [1] or chaperone proteins [2]. However, for the majority of proteins this necessity has not been documented [3] and many proteins also *in vivo* fold without intervention of any cellular supporting systems [4]. The basic paradigm that the entire information necessary for proper folding of a protein is contained in its sequence remains true. Thus the studies on protein folding *in vitro* provide an important step towards the understanding of protein folding *in vivo*.

The studies on protein folding have also a more practical aspect. Protein structures preceding the native form during folding, the so

*Presented at the symposium "Conformation of peptides, proteins and nucleic acids" held in Rynia, Poland, on 26–29 May, 1999.

✉Tel: (48 22) 659 7072 ext. 2351; fax: (48) 3912 1623; e-mail: michald@ibb.waw.pl

Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; TFE, trifluoroethanol.

called folding intermediates, are the most important structures responsible for conformational diseases of proteins caused by aggregation of proteins [5, 6]. Some serious diseases of humans and animals, like Alzheimer disease, Parkinson disease, Huntington disease, cystic fibrosis or spongiform encephalopathies are among them. The list of protein conformational diseases is expanding fast [7, 8]. The design of potential therapeutics requires the knowledge of their targets, in this case the protein folding intermediates. Also, the potential for biological activity of partly folded proteins like prions [9] or some proteins of the signal transduction cascade [10, 11] has been revealed recently. Genomics delivers protein sequences, but only a profound knowledge of the rules of structure formation in proteins can lead to structural proteomics, i.e. cataloguing of protein structures and functions in living organisms. Biotechnology also can benefit from protein folding studies since the designing of new proteins requires that they are properly folded.

The organisation of protein structures is in general based on the intertwining elements of secondary structure like α-helices, β-strands and turns [12]. These secondary structural elements are stabilised by local interactions, acting within a given element, but also by long range interactions, between residues participating in different structural elements. The long range interactions establish the tertiary structure of a protein. The aim of protein folding studies is to elucidate the location, mechanism, and time course of stabilisation of these basic structure-building blocks. In spite of a significant effort, the way by which the structure in proteins is formed has not been elucidated yet, and controversies are abundant in the literature. Available data, however, make it clear that the basic shape of the protein, in other words the general topology of protein chain, is established early in the folding process. This review focuses on questions and answers dealing with the early steps of folding.

Most important questions can be put as follows. What structural elements can be detected in the unfolded state? In other words one needs to know what elements of secondary structures (native or not) occur in unfolded proteins and what are their stabilities? What interactions stabilise these, so called, residual structures? From this the next question follows: do these residual structures help at all in folding of proteins, i.e. do they accelerate the folding process by decreasing the height of the barrier of the folding transition state? Are the local interactions sufficient to drive proper folding, or these local events are irrelevant, and long-range, strong, hydrophobic interactions are those which fold proteins? Different research groups gave different answers to these questions, and though the definitive answers are still to come, recent research provided important information. Theoretical studies also address the above questions and thorough reviews describe the development of the theoretical approach [13–18]. Only experimental results are discussed here.

Due to cooperativity of folding the structures attained during different stages of folding cannot be studied at equilibrium. The studies of the early steps of folding also suffer from this difficulty and the conformational state of the protein at the beginning of folding is usually modelled by short peptides, protein fragments, truncated proteins, or proteins denatured by some agent, e.g. a chaotropic agent. All these approaches have been undertaken to address the first question on the structural nature of the unfolded protein, when the folding process is triggered by applying conditions favouring folding. The following sections present first the studies of the structure of isolated short peptides, natural, modified or designed, protein fragments, truncated proteins, and finally of the entire denatured proteins both in unfolding and folding conditions.

SHORT PEPTIDES

Short peptides provide a perfect model system for studying local interactions, as all elements of secondary structure like α-helices, fragments of β-strands named β-hairpins, and β-turns can be quite stable in peptides. Initial doubts as to the potential stability of the α-helix in short peptides have been dispelled by the discovery [19, 20] that the N-terminal fragment of RNase A, consisting of twelve amino-acid residues, is partly helical in physiological conditions. Intensive studies followed, and soon peptides of remarkable helical content have been found: the modified C-peptide of RNase A [21] is in 50% helical; H-helix of myoglobin is in 30% helical [22], λ-repressor fragment – 40% helical [23], C-terminal peptide of Rnase H is 40% helical [24], the fragment of ribosomal protein L9 is in 53% helical [25]. The factors stabilising the α-helical structure in peptides have already been thoroughly studied and several reviews summarise the obtained data [26–28]. The collected data allowed to design algorithms that can in many cases properly estimate the stability of helical structure in a given peptide [29].

The most general conclusion drawn from these studies is that the backbone itself has an inherent tendency to form a helix [30]. In this work, it is suggested that a β-carbon, which is present in almost all amino-acids, should be treated as part of the backbone, thus poly-alanine is used as the model of the polypeptide backbone, and the studies of a series of alanine based peptides provided most precise information on the factors influencing the stability of α-helical structure [31]. In general, replacement of Ala with amino acids which introduce more flexibility (Ala to Gly) destabilises the helix by stabilising the unfolded state entropically. Bulky side chains also destabilise the helix entropically because the helical backbone sterically decreases the conformational space accessible for the side chain [32]. The amino acids with side chains destabilising helical conformation to the smallest extent are called helix formers. Numerous studies allowed to establish the helical propensities of all amino acids [32–35].

On the other hand, specific interactions between side chains may stabilise the helix even by more than 1 kcal/mole per interaction. A wide spectrum of different pairwise interactions has been catalogued, like salt bridges [36], interactions between non-polar groups [37–38]; charged and aromatic groups [39, 40] or hydrogen bonds [41]. The sign and magnitude of these interactions are strongly dependent on mutual arrangement of the interacting groups on the helix cylinder. Also, interactions of N- and C-terminal residues with backbone atoms, so called N-cap or C-cap interactions, respectively, may stabilise the helix. Simultaneously, these interactions can stop further helix propagation providing helix termination signals for helices in proteins. Therefore they may play an important role in the formation of general protein chain topology [42], although they not always prevent helix propagation [43]. The data on interactions providing helix termination signals have been summarised recently [44].

Formation of an α-helix is a fast process. Kinetic studies estimate the rate of helix elongation to be $10^8 \times s^{-1}$. However, it is the formation of the first helical turn, the so called nucleation step, which is rate limiting in helix formation, thus helix formation times are in the range of a microsecond [45–47].

Calorimetric studies of the helical stability of an alanine based peptide have shown that an α-helix is stabilised enthalpically [48]. Hydrogen bonds, each stabilising an α-helix by 1 kcal/mole have been suggested as a factor responsible for this enthalpic stabilisation. The enthalpic stabilisation of helices seems to be stronger in isolated peptides than in helices buried in proteins [49]. It has been suggested that, in alanine based peptides, the backbone peptide groups are not dehydrated and the dehydration following the burial of an α-helix in proteins leads to the decrease of the strength of helix stabilising hydrogen bonds

[42, 50, 51]. Dehydration of backbone atoms, which destabilises folded forms, has been suggested to explain the discrepancies in thermodynamic data obtained for folding of proteins [52]. Chaotropic agents destabilise helices [53] and some alcohols like trifluoroethanol (TFE) stabilise their structure [54]. The mechanism of stabilisation by TFE is a matter of discussion. Usually it is assumed that TFE stabilises the hydrogen bonds. However, recent experiments do not support this view [55]. As they suggest that the presence of alcohol shifts the equilibrium towards the helical structure not by influencing the free energy of the helix itself but by increasing the free energy of the unfolded state by disfavouring hydration of the backbone atoms. Although the correlation between the amino-acid type and helix formation tendency has been well established the knowledge on the nature of forces responsible for helix stability is not sufficient yet.

The studies of fragments of β-structures in isolated peptides, named β-hairpins, which involve only two antiparallel strands, are more difficult than those of α-helices because of an inherent tendency of such structures to aggregate. Most of the studied peptides encompassing β-hairpins in proteins are either unstructured or aggregate [56]. Nevertheless, in recent years several peptides have been found in which a significant population of the β-structure has been detected in monomeric species [56–64]. The shortest of these peptides consists of 8 amino acids [62]. Interestingly, in natural peptides β-hairpins of significant stability are a rare event, more often they are detected in modified or *de novo* designed peptides.

It has been suggested that β-hairpins are stabilised by a network of different interactions of the side chains and backbone atoms [63, 64], including also electrostatic interactions [65]. A hydrophobic interaction involving tryptophan residue is postulated as the stabilising factor for a β-hairpin in a model peptide [63]. ^{13}C-NMR relaxation studies show that, upon formation of a β-hairpin in a model peptide, the mobility both of the backbone and side-chain atoms becomes restricted [66, 67]. The kinetics of folding of a β-hairpin peptide has been measured by the temperature-jump method [63]. It has been shown that folding of a β-hairpin is about 30-fold slower than folding of an α-helix. Two factors can contribute to the observed differences; first, in a propagating structure a single H-bond in β-hairpin fixes two residues as compared to one in an α-helix, and second, an α-helix formation can be started at any residue and in the case of a β-hairpin only at turn residues [63].

Similarly to α-helix, β-hairpin formation is enthalpy driven, H-bonds being the most probable candidate for the enthalpic component [63, 64]. Also, like in the case of α-helix chaotropic agents unfold β-hairpins [68] and TFE stabilises its structure [62]. Different approaches have been used to study the β-sheet preferences of different amino acids. Based on database structure analysis [69], comparison of the stability of protein mutants [70–72] and model peptide studies [73, 74] a series of preferences has been obtained. However, the generality of the observed preference scale has been questioned [75, 76]. In a *de novo* designed β-hairpin the replacement of strand residues by alanines leads to the destabilisation of the hairpin structure [77]. This has been interpreted as indicating the side-chain-side-chain interactions. However, the side chains may exert their effect in an indirect way [64], for instance by modulating the intrinsic propensities or strength of hydrogen bonds [78], and not necessarily by a direct interaction.

Turn-like structures have also been detected in short peptides [79–82]. The turn residues may enforce the type of the β-hairpin formed [77, 83, 84]. Proper turn formation has been shown to be required for proper folding of P22 tailspike protein [85]. Amino-acid preferences to form a turn have been measured [86].

FACTORS DIFFERENTIATING THE α-HELICES AND β-STRUCTURES

As protein structure architecture is built upon this differentiation the question of factors responsible for formation of helices in some parts of protein sequence and sheets in the others is a fundamental one. Remarkably, thermodynamic studies point to H-bonds as the main stabilising factor of both helices and hairpins in isolated peptides. However, H-bonds are believed to be unspecific [87] and thus unable to provide necessary differentiation. Similarly, hydrophobic interactions shown in some cases to stabilise local structures, are non-specific. Thus the main driving factor for formation of the α-helices or β-hairpins seems to be the stereochemical code of intrinsic propensities which disfavours helical conformation for some amino acids and sheet conformation for the others [38, 88]. This code may be modulated by neighbouring residues [89]. Also, external conditions, like trifluoroethanol, may change the rank order of the hierarchy of preferences of particular amino acids [51].

In the case of α-helix the stereochemical code of conformational preferences results from the fact that the helix backbone imposes conformational restrictions on side chains leading to large entropy loss upon helix formation. It has been pointed out [42] that the destabilisation of a helical conformation in itself might be sufficient to shift the equilibrium towards a β-hairpin without any special β-sheet stabilising factors. In this model, the molecules for which population of the helical Φ, Ψ, minimum of a Ramachandran plot is disfavoured, would tend to populate the second minimum which corresponds to a β-hairpin. Random coil might be dominating in peptides of mixed tendencies. An interesting alternative has been suggested based on the correlation between the β-sheet forming tendencies and the hydrogen exchange rates of amide protons of different amino acids [78]. Namely, it has been suggested that the stability of a hydrogen bond may depend on the side-chain and its conformation. In any case, the side-chains stabilise a given structure in an indirect way, by affecting the properties of the main chain, and not necessarily by a direct side-chain-side-chain interaction [64]. The role of the backbone in protein folding has only recently been recognised [52]. In 1996 Honig and Cohen entitled their paper: "Adding backbone to protein folding: proteins are polypeptides" [90]. In a series of studies it has been found that the scales for secondary structure propensities obtained in peptide studies correlate well with the statistical propensities for populating specific regions of the Φ, Ψ space in protein crystal structures [69, 91].

FRAGMENTS OF PROTEINS

The search for folding initiation sites has also been carried out in the fragments of proteins [92, 93]. The list of proteins for which the conformational analysis of their fragments has been carried out is long. Among them are the studies of plastocyanin [94], lysozyme [95, 96], barnase [97], BPTI [98], cytochrome *c* [99], chymotrypsin inhibitor 2 [100], protein GB1 [101], myoglobin [102], ribosomal protein L9 domain [25] or a series of three different proteins [103]. In the case of α-spectrin SH3 domain fragments some non-native structures have been found [104]. In all these studies it has been shown that, in general, the stability of the detected residual structures is weak, although a deviation from random coil seems to be a common feature. Stable local structures in protein fragments are rather an exception than a rule.

Similar results have been obtained in the studies of truncated proteins in which several terminal residues have been removed leading to the unfolding of such proteins as staphylococcal nuclease [105–107], cytochrome *c* [108], or chymotrypsin inhibitor 2 – CI2 [109]. In CI2 the removal of C-terminal amino acids results in the disordered protein, in

which, however, hydrophobic patches could be detected.

In conclusion, in the isolated fragments of proteins the elements of structure can be detected, usually of low stability, although it is relatively easy to obtain their significant stabilisation by rational modifications. This observation has led to the suggestion that the sequences of protein were evolutionarily selected for the lack of stable structures in isolated fragments [87] and that this instability may be necessary for optimisation of folding [110]. The relative instability of local structures may be helpful at later steps of folding and facilitate the subtle changes in local structures necessary to obtain the final, close, tertiary contacts.

RESIDUAL STRUCTURES IN DENATURED PROTEINS

In parallel to the conformational studies of protein fragments, studies were carried out on entire denatured proteins. The residual structures are even more difficult to detect in the entire sequences than in isolated fragments, since they span only a fragment of the sequence, and are populated only in a fraction of time. The signals coming from unfolded parts of a molecule in classical methods of structure analysis, like circular dichroism or nuclear magnetic resonance, may interfere with detection of weakly stable structures. Numerous studies on the structure of unfolded proteins in unfolding external conditions have been reported [111–118]. In general, these experiments did not detect stable structures in fragments of protein sequences but detected signals of existing non-random, fluctuating structures. The development of NMR spectroscopy allowed to obtain peak assignments for at least ten unfolded proteins [119]. In myoglobin [120] the fluctuating, weak structural elements are detectable independently of the degree of protein unfolding. NMR studies detected local hydrophobic interactions [112], local restrictions of main chain mobility [113] and nascent turns [121].

UNFOLDED PROTEIN IN FOLDING CONDITIONS

Upon transition to folding conditions the denatured protein chains undergo contraction in the so called burst phase. This phase is usually completed within the dead time of the experimental instrumentation used in the kinetic studies of folding. The contracted burst-phase state is thought to be formed in the microsecond time scale. According to some authors it is the result of a non-specific hydrophobic collapse and thus without relevance to specific structure formation [122–124]. Data collected by others point to the contrary, indicating the presence of the elements of native-like structures even during the earliest folding events (for a review see [125]).

In some cases the unfolded protein chain can be maintained and studied at equilibrium in folding conditions. For instance, some disulphide bonded proteins unfold when the disulphide bonds are reduced. By applying reducing conditions the unfolded proteins can be studied in folding conditions. In the most thoroughly studied case of bovine pancreatic trypsin inhibitor (BPTI) native-like tendencies have been detected and shown to accelerate native pairing of one of the disulphide bonds [126, 127]. The residual structure, which is responsible for the above effect, is stabilised by an interaction of the two stretches of hydrophobic residues [128]. More detailed studies [64] showed that the residual structure, most probably a native-like β-hairpin, is weakly populated (below 10%) in reduced BPTI. However, its presence accelerates the overall formation of the natively paired disulphide. These data provide an experimental evidence that, in unfolded proteins in folding conditions, native-like tendencies, though thermodynamically not very strong, can be detected and can affect the folding process. The presence of the

native-like structure in reduced BPTI has been confirmed by detection of a specific binding of this molecule to target enzyme [129]. It has been shown previously [130] that, if some structure accelerates the formation of a disulphide, the same structure becomes stabilised after the disulphide is formed. And indeed, NMR studies of a BPTI variant have shown that the formation of the disulphide stabilises the native β-hairpin [131, 132]. Fluorescence transfer studies also suggested the presence of a residual structure in reduced BPTI [133–135]. This structure escaped earlier detection due to inability of classical methods of structural analysis to detect structures populated at the level of a few percent. Disulphide bonds were also used as reporter groups to show that a molten globule of α-lactalbumin has a native-like fold [136].

In some dimeric proteins, like Arc repressor, folding is coupled to the formation of dimers. This allowed to study the unfolded Arc repressor structure in folding conditions. Monomers were shown to be partly structured [137].

BURST PHASE STRUCTURES

Kinetic studies of early folding events provide some insight into the structures appearing in the burst phase. Some of these tests point to lack of secondary structure [108, 122–124, 138]. It has been suggested that, in cytochrome *c*, folding in the early phase is a time consuming process of energetically uphill barrier crossing and not the downhill collapse [122]. These data suggest that specific folding starts from a slow process of formation of a nucleus stabilised by long range interactions of hydrophobic residues [124].

Other data point to the contrary [125]. The burst phase structures usually give rise to a significant fraction of a native-like circular dichroism signal [139]. Some protection of amide hydrogens against exchange has also been observed in the earliest detectable folding events in numerous kinetic studies of folding [125]. Stabilisation of early folding β-hairpins by clusters of hydrophobic residues in interleukin-1β has been demonstrated by an NMR spectroscopic study [140]. In the immunoglobulin fold a β-strand conformation is established early in folding [141]. Specific changes of tryptophan fluorescence develop in the burst phase in ubiquitin [142, 143] indicating rather specific than non-specific tryptophan interactions. Single mutations were observed to significantly destabilise an early folding intermediate in ubiquitin [143] and cytochrome *c* [144] providing evidence for non-random interactions. It has also been suggested that in some conditions (like low temperature) local interactions prevail whereas in the other long-range interactions become stronger [145].

In some cases a non-native secondary structure develops in the burst phase. The best known example is β-lactoglobulin [146], where the CD signal indicating the presence of an α-helix, which develops in the burst phase, disappears later during folding. The SH3 domain engineered to exhibit non-native α-helix propensities, and containing non-native helix in the denatured state [147, 148] finally folds to its native structure. Improper local structures do not prevent a protein from folding. In some cases, the context masks these local structural propensities [76, 86, 149]. That has been clearly shown in a "chameleon" sequence [149] where an eleven residue peptide becomes either helical or forms a β-sheet in the folded protein depending on its position in the protein sequence. By no means these data provide evidence for a general mechanism of breaking down the local preferences in the later steps of folding. They point out that in some cases non-local forces can overcome non-native propensities at the possible cost of the efficiency of the folding process.

FOLDING INITIATION SITES

Residual structural elements in unfolded proteins have been postulated to provide folding initiation sites [94, 100, 150-152] and, more specifically, β-hairpins were suggested to serve as folding initiation structures [153]. Residual structures are expected to accelerate folding by narrowing the accessible conformational search space. On the other hand, it has been argued that the marginal stability of local structures lowers their population to the levels of at best a few percent of the total, and thus role of these structures may be marginal in comparison to much stronger, long range hydrophobic interactions [154].

The influence of local structures appearing in unfolded state on the rate of folding is a matter of vigorous discussion. In various phenomenological models of the folding process elements of secondary structure play different roles. In the framework model [155, 156] local secondary structure elements are crucial. This model assumed the formation of stable elements of secondary structure at the early phase of folding and their subsequent simple recombination. In the "puzzle" model [157] folding was thought to start at any region of sequence in the protein, so that no specific folding initiation sites seemed necessary. The diffusion-collision model [158] assumed the stabilisation of the pre-formed local structures by collisions at a rate limited by the diffusion rate. The nucleation model [159] assumed a single, stable nucleation structure and further growth of the rest of structure on the template of the nucleus. The nucleation-condensation model [110] underscored the necessity for a low stability and flexibility of folding initiation sites (named nucleation sites). Hydrophobic collapse models emphasise the role of the long range hydrophobic interactions and postulate that local structures are irrelevant for folding [160].

UNSTABLE FOLDING INITIATION SITES CAN ACCELERATE FOLDING

It is mostly agreed that, in the burst phase, a state of pre-equilibrium of different conformations is established in the very fast process deprived of energetic barriers – energetically "downhill". This pre-equilibrium precedes the transition of a free energy barrier what leads to the native state [125]. The presence of this barrier has been postulated to be a general phenomenon in protein folding, because protein folding kinetics are in general monoexponential. The transition state theory is widely applied for studying this barrier, although the validity of this approach is still under discussion. By analogy with the transition state of an ordinary chemical reaction this barrier is called the transition state barrier for folding. Recent studies, however, indicate that the character of this energetic barrier for folding is in many respects different from the transition state barrier of an ordinary chemical reaction.

An exception in the character of an early folding step has been suggested in the case of cytochrome *c* in which the entire folding process is thought to proceed uphill of the folding barrier [122] without a specific pre-equilibrium state. This conclusion came from the conviction that the burst phase in the case of cytochrome *c* is a non-specific hydrophobic contraction in response to changed external conditions, and thus is not treated as part of the folding process [124]. However, the entire folding process is the response to the changed external conditions, the burst phase included.

The rate limiting step of folding is the crossing of the transition state barrier, as the fraction of molecules populating the high energy transition state is limited by the height of the corresponding energy barrier. Whether the rate limiting step is due to water exclusion or internal friction of side chain rearrangements

still remains a matter of debate [161]. Since the rate limiting step seems to be kinetically separated from the burst phase by the state of pre-equilibrium, the role of local structures on overall folding is sometimes questioned [160, 162].

The following simple reasoning shows that even weakly stable local structures can accelerate the formation of long range contacts necessary for crossing the transition state barrier, and can thus accelerate the entire folding process. A simple scheme describing formation of a long range contact in an ensemble of pre-equilibrated weak structures is considered (Fig. 1). The model assumes that a spectrum of short range local interactions (AxB) exists in a state of pre-equilibrium with a fast interconversion between a set of unfolded forms (u ensemble) and a local native-like structure (c ensemble). The native-like local structure may entropically accelerate the formation of the long range contact (CxD) in the c subset of molecules, as it decreases the conformational search space in the c molecules. The observed rate of formation of the long-range contact for the entire population of the molecules (k_{obs}) depends on the rates of long range contact formation in u and c ensembles and the pre-equilibrium constant (K). For simplicity the unfolding rates have not been taken into account.

$$k_{obs} = k_u * f_u + k_c * f_c ;$$

where f_c and f_u are the fractions of c and u, respectively. As c and u are assumed to equilibrate fast $f_c/f_u = K$, $f_c = K/(1+K)$ and $f_u = 1/(1+K)$ are constant, time independent, values, and

$$k_{obs} = (k_u + k_c * K)/(1+K) \text{ or}$$
$$k_{obs} = k_u * (1 + K*k_c/k_u)/(1+K)$$

The observed rate enhancement, caused by the structure in subset c of the molecules, is $r_{obs} = k_{obs}/k_u$ and:

$$r_{obs} = (1 + K * r_{max})/(1+K) \qquad (1)$$

where r_{max} is the maximum rate enhancement possible $r_{max} = k_c/k_u$.

The dependence of the observed rate enhancement for the entire population of the molecules (r_{obs}) on the pre-equilibrium constant K for a series of different values of r_{max} is given in Fig. 1b.

It is clear that a significant acceleration of k_{obs} is possible even for quite low values of K, below 0.1, when the population of a folding initiation site is at the level of a few percent or lower. It is impossible to detect these structures by classic methods of structural analysis. However, these weak structures can significantly accelerate acquisition of proper long-range contacts, provided they equilibrate fast with the remaining molecules and that their existence in a single molecule strongly influences the rate of formation of a long-range contact (r_{max} high). Fast pre-equilibrium ensures a constant flux towards c subpopulation, so that at any moment the same fraction of molecules occupies in the fast-folding c state. That is why the existence of a small c subpopulation can influence the folding kinetics of the entire population.

In folded proteins the effective concentration, reflecting the rate of formation of the contact between the two residues [127], even for long range contacts, can be very high, at the level of 10^7 moles [163]. In unfolded proteins the effective concentration is at the level of a few millimoles [127]. The values of effective concentration and thus the values of r_{max} of 10^3 in partly folded proteins are not unexpected. In the case of $K<<1$ the observed rate depends on the residual structure population in a linear fashion, provided in c molecules a long range contact formation is significantly accelerated. In other words, it is necessary that local structures increase considerably the effective concentrations of a potential long-range contact. Stable, fixed elements of secondary structures may not be the optimal in-

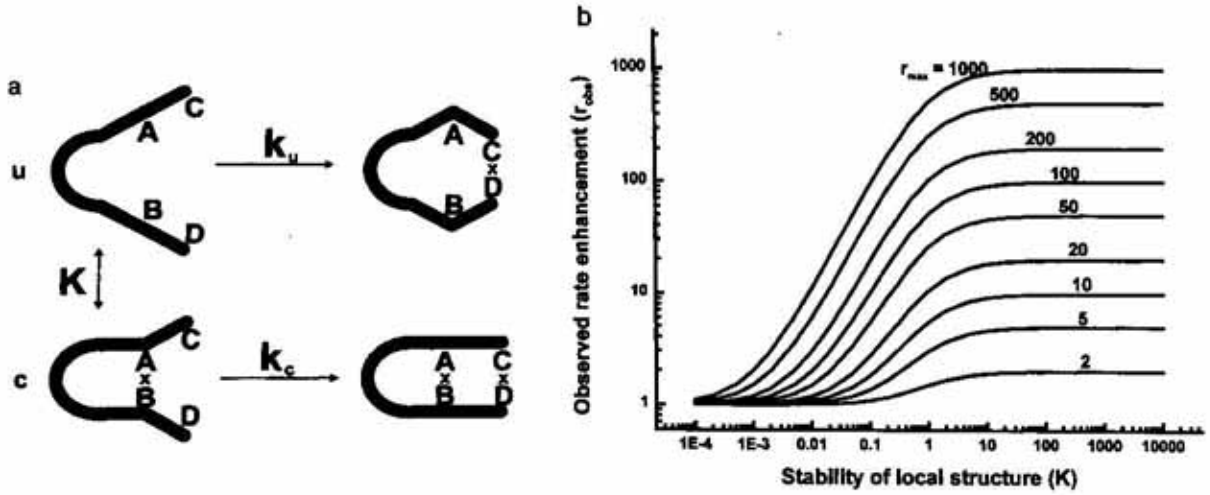

Figure 1. The ability of a local structure (AxB) of a low population to affect the rate of formation of a long range contact (CxD) in the entire population of folding molecules.

a. Scheme of four species populated in the early steps of folding.

The local interaction AxB present in the subset c of the molecules is in fast equilibrium with form u in which AxB do not interact. The equilibrium constant K describes the relative populations of c and u forms. In subset c the rate of formation of the long range contact (k_c) is different than in u (k_u) for entropic reasons. The overall rate of formation of a long range contact (k_{obs}) depends on K and the ratio of k_c and k_u ($r_{max} = k_c/k_u$), as described in the text.

b. The dependence of the observed rate enhancement ($r_{obs} = k_{obs} / k_u$) caused by the presence of the local structure of different stability (K) for several values of r_{max} calculated according to equation 1.

For higher r_{max} values even a small fraction of molecules containing AxB can affect the overall rate of formation of the long range contact. On the other hand, a further increase of K value above the level of 10 has no effect on folding kinetics.

termediate structures to fulfil this condition. The fact that local structures, even populated as only a small fraction of molecules, can influence the folding flux of the entire population of molecules, is often not sufficiently appreciated. Similar kinetic schemes seem to apply to the entire process of folding, where the overall folding rate may depend on the fraction of population of an early, burst-phase intermediate [125, 164].

RELEVANCE OF LOCAL STRUCTURES FOR FOLDING

The question at which stage of folding and by what interactions or mechanisms the topology of the main chain, i.e. the global structure of the protein, is established [42, 50], is still a matter of debate. Two different concepts were described in the literature. In the first, the final structure is being built in a hierarchic fashion, starting from local structures, which are then stabilised by long-range interactions. In the second, the entire folding is directed by hydrophobic collapse, and local structures, even if they exist, are of no relevance, as they do not accelerate folding. Both these models find support in theoretical work [165–168] and in experimental data [87]. Also some indirect arguments were proposed to support either possibility. For instance, it has been noted that hydrogen bonds in folded proteins are formed mainly between residues close in sequence, so that they are stabilised locally [169], and that helix termination signals are provided mainly by local interactions [44]. A correlation has been noted between the rate of folding and the fraction of local *vs.* long range contacts [170]. Proteins stabilised by a higher

fraction of local interactions fold faster. Recent experiments have shown that maintaining the hydrophobic core is not sufficient to retain the fold [171] and sometimes small changes are sufficient to change the protein fold [172] however, experiments with alanine replacements of a large portion of the sequence of Arc repressor point to the contrary [173]. Also, periodicity of hydrophobic-hydrophilic residues in protein sequence has been shown in several cases to be a sufficient factor to stabilise a specific structure [174, 175].

Many attempts to accelerate folding by stabilising local structures (or to decelerate it by destabilising them) have been undertaken, and some of them failed. In general, these experiments show that there is no simple dependence of the rate of folding on the stability of local structural elements.

The most straightforward case of the lack of dependence of the folding rate on the stability of constituent helices has been demonstrated in a simple system of the dimerisation of two α-helices into a coiled-coil structure [123]. In a coiled coil the two helices wind around each other, which leads to helix bending [176]. In the above experiment some residues outside of the contact interface were replaced by glycines, highly destabilising the helical structure of the monomers. However, the rate of formation of the dimer decreased only slightly. Interestingly, the decrease in the rates of unfolding was much more pronounced. This demonstrates that, in the case of a coiled coil, the best folding nucleation sites are not pre-formed helices but, most probably, nascent turns of helical conformation, which recombine with each other and zip up the leucine zipper structure. In agreement with this, in the transition state for coiled-coil formation a high population of intramolecular hydrogen bonds has not been detected [55]. However, it has to be taken into account that folding of a coiled coil has some features that may not be fully representative for general protein folding. First, in the coiled-coil the helices are distorted, thus regular helices are not necessarily the most energetically favourable templates for dimers (for a helical ensemble k_c may be lower than k_o, instead of being higher, using the terminology from Fig. 1). Second, the rate of the formation of a dimer is much higher than the folding of even small proteins, which indicates that the transition state barrier must be very low in this case. Anyway, the coiled-coil folding experiment provides support for the nucleation mechanism, with flexible structure formation as the starting point for folding. Similarly, several other studies did not detect any increase in folding rates upon stabilisation of local structures [177, 178].

In myoglobin the kinetics of folding has been compared in a series of variants with strongly destabilised helix H [178]. The only observable local structure in the fragments of myoglobin can be detected in helix H, which participates in the burst phase of folding and was suggested to be a folding initiation site for myoglobin [179]. It was shown that destabilising mutations decelerate folding, but only in the presence of more than 1 M urea, when the molten globule intermediate becomes destabilised. As shown in Fig. 1b, high stability of a folding nucleus (molten globule) decouples the overall folding rate from the local structure formation in helix H. In the absence of urea the nucleus is stable enough even with destabilised helix H. This is in agreement with the robustness of the folding code postulated recently [42, 50]. Robustness assumes that only a small fraction of early phase interactions is sufficient to induce proper folding. Similarly, the analysis of folding of a series of permutants of SH3 domain [180], which differ only by the region of protein sequence in which the N- and C-termini are placed, with different elements of local structures broken or stabilised, shows that the folding always leads to the same final native structure. None of the elements of secondary structure is indispensable for folding, as other elements seem always to be able to substitute it during

folding, which again points to the robustness of the folding code.

Other studies, however, reported a straightforward correlation between local structure stability and folding rates. Mutations of exposed polar residues enhancing native-like tendencies in helical segments of proteins were found to accelerate folding [181–183]. Fluoroalcohols which stabilise local interactions while destabilising hydrophobic forces have been found to dramatically accelerate folding of a small α/β protein [184]. The residual structure in unfolded BPTI, populated by less than 10%, can nevertheless accelerate several-fold native pairing of the native disulphide in the entire population of molecules [64]. Stabilisation of the native-like β-hairpin was suggested to be responsible for this effect. Also, it has been shown directly that the diffusion of the elements of secondary structure has a direct influence on the height of the main transition state barrier and thus on the rate of folding [185]. It has been found that, in barnase, local structures are strengthened in early folding steps by non-native tertiary interactions which are subsequently changed to native ones [186]. The effects of local interactions on protein stability have been reviewed by Munoz *et al.* [177].

Mutations in turn residues do not provide simple answers either. In some cases these substitutions lead to changes in the rate of folding [86] and in other they are irrelevant [187]. Selection of folded proteins from protein libraries randomised at turn positions have shown that sequences with proper turn residues are preferred thermodynamically [188]. In terms of protein structure some proteins tolerate mutations in turn regions [187, 189] and some don't [190], but usually these mutations affect stability [86,188] and folding kinetics.

In summary, the recently obtained data indicate that local, weak structures provide efficient folding initiation sites. Several interesting experiments, two of which have been discussed above, show that the role of these structures does not follow simple rules, and the results were interpreted sometimes as contradicting the mentioned above statement. However, as pointed out previously, different interpretations are possible, and this can explain the contradiction. More work is necessary to provide the final answer.

It is clear, however, that local interactions must stabilise the folding transition state to accelerate folding [191]. For instance, in the transition state for bimolecular folding of the S-protein-S-peptide complex, both the α-helical structure of the peptide and the hydrophobic contact are present [192]. The balance between local and long-range interactions seems to be maintained for optimisation of folding [87]. The long range, hydrophobic forces are very strong and can easily overcome weak forces acting locally. This does not seem to necessarily happen in protein folding. Rather, both factors act in parallel. The role of local structures is to facilitate and accelerate the formation of stabilising long range contacts in the subset of folding molecules. These long range contacts do not have to be native to fulfil their role [186]. Their formation may lead to a decrease of the transition state barrier in a given subset of the molecules and channel the entire population across the barrier [64], because of the pre-equilibrium existing in the unfolded state in folding conditions [193]. On the other hand overly stable, fixed elements of secondary structure do not seem to facilitate proper long range contacts. The data available underscore the gradual acquisition of fixed structure in the process of folding. It is also possible that the stability of local structures is maintained on the proper (low) level to facilitate the fit of local elements into the final structures without energy loss and excess barriers in the final step of side-chain fixation and water expulsion. The final protein structure, the physical basis for the entire biological activity of proteins, seems to be born in a process of mutual stabilisation of local and long-range contacts. During evolution, such protein sequences might have been selected in

which both types of interaction are precisely correlated to obtain foldable proteins.

REFERENCES

1. Hardesty, B., Tsalkova, T. & Kramer, G. (1999) Co-translational folding. *Curr. Opin. Struct. Biol.* **9**, 111–114.
2. Ellis, R.J. & Hemmingsen, S.M. (1989) Molecular chaperones: Proteins essential for the biogenesis of some macromolecular structures. *Trends Biochem. Sci.* **14**, 339–342.
3. Radford, S.E. (1998) What's new in protein folding? *Folding & Design* **3**, R59–R63.
4. Thulasiraman, V., Yang, C.-F. & Frydman, J. (1999) *In vivo* newly translated polypeptides are sequestered in a protected folding environment. *EMBO J.* **18**, 85–95.
5. Sifers, R.N. (1995) Defective protein folding as a cause of disease. *Nature Struct. Biol.* **2**, 355–357.
6. Wetzel, R. (1996) For protein misassembly this is the I decade. *Cell* **86**, 699–701.
7. Carrell, R.W. & Lomas, D.A. (1997) Conformational disease. *Lancet* **350**, 134–138.
8. Baum, J. & Brodsky, B. (1999) Folding of peptide models of collagen and misfolding in disease. *Curr. Opin. Struct. Biol.* **9**, 122–128.
9. Harrison, P.M., Bamborough, P., Dagget, V., Prusiner, S.B. & Cohen, F. (1997) The prion folding problem. *Curr. Opin. Struct. Biol.* **7**, 53–59.
10. Daughdrill, G.W., Chadsey, M.S., Karlinsey, J.E., Hughes, K.T. & Dahlquist, F.W. (1997) The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28. *Nature Struct. Biol.* **4**, 285–291.
11. Kriwacki, R., Hengst, L., Tennant, L., Reed, S.I. & Wright, P.E. (1996) Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11504–11509.
12. Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B. & Thornton, J.M. (1997) CATH – a hierarchic classification of protein domain structures. *Structure* **5**, 1093–1098.
13. Bryngelson, J.D., Onuchic, J.N., Socci, N.D. & Wolynes, P.G. (1995) Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Struct. Funct. Genet.* **21**, 167–195.
14. Karplus, M. & Sali, A. (1995) Theoretical studies of protein folding and unfolding. *Curr. Opin. Struct. Biol.* **5**, 58–73.
15. Dill, K.A. & Chan, H.S. (1995) Principles of protein folding. A perspective from simple exact models. *Protein Sci.* **4**, 561–602.
16. Dill, K.A. & Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19.
17. Shakhnovich, E. (1997) Theoretical studies of protein-folding thermodynamics and kinetics. *Curr. Opin. Struct. Biol.* **7**, 29–40.
18. Dobson, C.M., Sali, A., Karplus, M. (1998) Protein folding: A perspective from theory and experiment. *Angew. Chem. Int. Ed.* **7**, 29–40.
19. Brown, J.E. & Klee, W.A. (1971) Helix-coil transition of the isolated terminus of ribonuclease. *Biochemistry* **10**, 470–476.
20. Bierzynski, A., Kim, P.S. & Baldwin, R.L. (1982) A salt bridge stabilizes the helix formed by isolated C-peptide of RNase A. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2470–2474.
21. Osterhout, J., Jr., Baldwin, R.L., York, E.J., Stewart, J.M., Dyson, H.J. & Wright, P.E. (1989) 1H NMR studies of the solution conformations of an analogue of the C-peptide of ribonuclease A. *Biochemistry* **28**, 7059–7064.

22. Waltho, J.P., Feher, V.A., Merutka, G., Dyson, H.J. & Wright, P.E. (1993) Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin. *Biochemistry* **32**, 6337–6347.

23. Marqusee, S. & Sauer, R.T. (1994) Contributions of a hydrogen bond/salt bridge network to the stability of secondary and tertiary structure in lambda repressor. *Protein Sci.* **12**, 2217–2225.

24. Goedken, E.R., Raschke, T.M. & Marqusee, S. (1997) Importance of the C-terminal helix to the stability and enzymatic activity of *Escherichia coli* ribonuclease H. *Biochemistry* **36**, 7256–7263.

25. Luisi, D.L., Wu, W.J. & Raleigh, D.P. (1999) Conformational analysis of a set of peptides corresponding to the entire primary sequence of the N-terminal domain of the ribosomal protein L9: Evidence for stable native-like secondary structure in the unfolded state. *J. Mol. Biol.* **287**, 395–407.

26. Baldwin, R.L. (1995) Alpha-helix formation by peptides of defined sequence. *Biophys. Chem.* **55**, 127–135.

27. Chakrabartty, A. & Baldwin, R.L. (1995) α-Helix stability *Adv. Protein Chem.* **46**, 141–176.

28. Munoz, V. & Serrano, L. (1995) Helix design, prediction and stability. *Curr. Opin. Biotechnol.* **6**, 382–386.

29. Munoz, V. & Serrano, L. (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct. Biol.* **1**, 399–409.

30. Creamer, T.P. & Rose, G.D. (1992) Side-chain entropy opposes α-helix formation but rationalizes experimentally determined helix-forming propensities. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5937–5941.

31. Marqusee, S., Robbins, V.H. & Baldwin, R.L. (1989) Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5286–5290.

32. Creamer, T.P. & Rose, G.D. (1994) α-Helix forming propensities in peptides and proteins. *Proteins: Struct. Funct. Genet.* **19**, 85–97.

33. Chakrabartty, A., Kortemme, T. & Baldwin, R.L. (1994) Helix propensities of the amino-acids measured in alanine-based peptides without helix-stabilizing side chain interactions. *Protein Sci.* **3**, 843–852.

34. Blaber, M., Zang, X. & Matthews, B. (1993) Structural basis of amino acid α-helix propensity. *Science* **260**, 1637–1640.

35. O'Neill, K.T. & DeGrado, W.F. (1990) A thermodynamic scale for the helix forming tendencies of the commonly occurring amino acids. *Science* **250**, 646–650.

36. Huyghues-Despointes, B.M., Scholtz, J.M. & Baldwin, R.L. (1993) Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci.* **2**, 80–85.

37. Padmanabhan, S. & Baldwin, R.L. (1994) Tests for helix-stabilizing interactions between various nonpolar side chains in alanine-based peptides. *Protein Sci.* **3**, 1992–1997.

38. Creamer, T.P. & Rose, G.D. (1995) Interactions between hydrophobic side chains within alpha-helices. *Protein Sci.* **4**, 1305–1314.

39. Armstrong, K.M., Fairman, R. & Baldwin, R.L. (1993) The (i, i + 4) Phe-His interaction studied in an alanine-based alpha-helix. *J. Mol. Biol.* **230**, 284–291.

40. Dadlez, M., Bierzyński A., Godzik A., Sobocińska, M. & Kupryszewski, G. (1988) Conformational role of His-12 in C-peptide of ribonuclease A. *Biophys. Chem.* **31**, 175–181.

41. Huyghues-Despointes, B.M., Klingler, T.M. & Baldwin, R.L. (1995) Measuring the strength of side-chain hydrogen bonds in peptide helices: The Gln.Asp (i, i + 4) interaction. *Biochemistry* **34**, 13267–13271.

42. Baldwin, R.L. & Rose, D.G. (1999) Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem. Sci.* **24**, 26–33.

43. Jimenez, M.A., Munoz, V., Rico, M. & Serrano, L. (1994) Helix stop signals in proteins and peptides: The capping box does not necessarily prevent helix elongation. *J. Mol. Biol.* **242**, 487–496.

44. Aurora, R. & Rose, G.D. (1998) Helix capping. *Protein Sci.* **7**, 21–38.

45. Williams, S., Causgrove, T.P., Gilmanshin, R., Fang, K.S., Callender, R.H., Woodruff, W.H. & Dyer, R.B. (1996) Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* **35**, 691–697.

46. Thompson, P.A., Eaton, W.A. & Hofrichter, J. (1997) Laser temperature jump study of the helix-coil kinetics of an alanine peptide interpreted with a "kinetic zipper" model. *Biochemistry* **36**, 9200–9210.

47. Gilmanshin, R.S., Williams, R.H., Callender, W.H., Woodruff, W.H. & Dyer, R.B. (1997) Fast events in protein folding: Relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3709–3713.

48. Scholtz, J.M., Marqusee, S., Baldwin, R.L., York, E.J., Stewart, J.M. Santoro, M. & Bolen, D.W. (1991) Calorimetric determination of the enthalpy change for the α-helix to coil transition of an alanine peptide in water. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2854–2858.

49. Myers, J.K., Pace, C.N. & Scholtz, J.M. (1997) A direct comparison of helix propensity in proteins and peptides. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 2833–2837.

50. Baldwin, R.L. & Rose, D.G. (1999) Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem. Sci.* **24**, 77–83.

51. Luo, P. & Baldwin, R.L. (1999) Interaction between water and polar groups of the helix backbone: An important determinant of helix propensities. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4930–4935.

52. Yang, A.-S., Sharp, K.A. & Honig, B. (1992) Analysis of the heat capacity dependence of protein folding. *J. Mol. Biol.* **227**, 889–900.

53. Scholtz, J.M., Barrick, D., York, E., Stewart, J.M. & Baldwin, R.L. (1995) Urea unfolding of peptide helices as a model for interpreting protein unfolding. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 185–189.

54. Luo, P. & Baldwin, R.L. (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **36**, 8413–8421.

55. Kentsis, A. & Sosnick, T.R. (1998) Trifluoroethanol promotes helix formation by destabilising backbone exposure; Desolvation rather than native hydrogen bonding defines the kinetic pathway of dimeric coiled coil folding. *Biochemistry* **37**, 14613–14622.

56. Blanco, F., Ramirez-Alvarado, M. & Serrano, L. (1998) Formation and stability of beta-hairpin structures in polypeptides. *Curr. Opin. Struct. Biol.* **8**, 107–111.

57. Blanco, F.J., Jimenez, M.A., Herranz, J., Rico, M., Santoro, J. & Nieto, J.L. (1993) NMR evidence of a short linear peptide that folds into a β-hairpin in aqueous solution. *J. Am. Chem. Soc.* **115**, 5887–5888.

58. Blanco, F.J., Rivas, G. & Serrano, L. (1994) A short linear peptide that folds into a native stable beta-hairpin in aqueous solution. *Nature Struct. Biol.* **1**, 584–590.

59. Searle, M.S., Williams, D.H. & Packman, L.C. (1995) A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non-native beta-hairpin. *Nature Struct. Biol.* **2**, 999–1006.

60. de Alba, E., Jimenez, M.A., Rico, M. & Nieto, J.L. (1996) Conformational investigation of designed short linear peptides able to fold into beta-hairpin structures in aqueous solution. *Folding & Design* **1**, 133–144.

61. Sieber, V. & Moe, G.R. (1996) Interactions contributing to the formation of a beta-hairpin-like structure in a small peptide. *Biochemistry* **35**, 181–188.

62. Ramirez-Alvarado, M., Blanco, F. & Serrano, L. (1996) De novo design and structural analysis of a model β-hairpin peptide system. *Nature Struct. Biol.* **3**, 604–611.

63. Munoz, V., Thompson, V.A., Hofrichter, J. & Eaton, W.A. (1997) Folding dynamics and mechanism of beta-hairpin formation. *Nature* **390**, 196–199.

64. Zdanowski, K. & Dadlez, M. (1999) Stability of the residual structure in unfolded BPTI in different conditions of temperature and solvent composition measured by disulphide kinetics and double mutant cycle analysis. *J. Mol. Biol.* **287**, 433–445.

65. Blasie, C.A. & Berg, J.M. (1997) Electrostatic interactions across a β-sheet. *Biochemistry* **36**, 6218–6222.

66. Friedrichs, M.S., Storch, T.R., Bruccoleri, R.E., Mueller, L. & Constantine, K.L. (1995) Structural and dynamic properties of a β-hairpin forming linear peptide. 2. ^{13}C relaxation analysis. *J. Am. Chem. Soc.* **117**, 10864–10885.

67. Ramirez-Alvarado, M., Daragan, V., Serrano, L. & Mayo, K.H. (1998) Motional dynamics of residues in a β-hairpin peptide measured by ^{13}C-NMR relaxation. *Protein Sci.* **7**, 720–729.

68. de Alba, E., Blanco, F.J., Jimenez, M.A., Rico, M. & Nieto, J.L. (1995) Interactions responsible for the pH dependence of the β-hairpin conformational population formed by a designed linear peptide. *Eur. J. Biochem.* **233**, 283–292.

69. Munoz, V. & Serrano, L. (1994) Intrinsic secondary structure propensities of the amino acids, using statistical Φ-Ψ matrices: Comparison with experimental scales. *Proteins: Struct. Funct. Genet.* **20**, 301–311.

70. Regan, L. (1994) Protein structure. Born to be beta. *Curr. Biol.* **4**, 656–658.

71. Minor, D.L. & Kim, P.S. (1994) Measurement of the β-sheet-forming propensities of amino acids. *Nature* **367**, 660–663.

72. Smith, C.K., Withka, J.M. & Regan, L. (1994) A thermodynamic scale for β-sheet forming tendencies of the amino-acids. *Biochemistry* **33**, 5510–5517.

73. Kim, C.A. & Berg, J.M. (1993) Thermodynamic β-sheet propensities measured using a zinc finger host peptide. *Nature* **362**, 267–270.

74. Serrano, L. (1995) Comparison between the Φ distribution of the amino acids in the protein database and NMR data indicates that amino acids have various Φ propensities in the random coil conformation. *J. Mol. Biol.* **254**, 322–333.

75. Otzen, D.E. & Fersht, A.R. (1995) Side-chain determinants of β-sheet stability. *Biochemistry* **34**, 5718–5724.

76. Minor, D.L. & Kim, P.S. (1994) Context is a major determinant of beta-sheet propensity. *Nature* **371**, 264–267.

77. Ramirez-Alvarado, M., Blanco, F.J., Niemann, H. & Serrano, L. (1997) Role of β-turn residues in β-hairpin formation and stability in designed peptides. *J. Mol. Biol.* **273**, 898–912.

78. Bai, Y. & Englander, S.W. (1994) Hydrogen bond strength and β-sheet propensities: The role of a side chain blocking effect. *Proteins: Struct. Funct. Genet.* **18**, 262–266.

79. Yao, J., Dyson, H.J. & Wright, P.E. (1994) Three-dimensional structure of a type VI turn in a linear peptide in water solution. Evidence

for stacking of aromatic rings as a major stabilising factor. *J. Mol. Biol.* **243**, 754–766.

80. Yao, J., Feher, V.A., Espejo, B.F., Reymond, M.T., Wright, P.E. & Dyson, H.J. (1994) Stabilisation of a type VI turn in a family of linear peptides in water solution. *J. Mol. Biol.* **243**, 736–753.

81. Dyson, H.J., Cross, K.J., Houghten, R.A., Wilson, I.A., Wright, P.E. & Lerner, R.A. (1985) The immunodominant site of a synthetic immunogen has a conformational preference in water for a type-II reverse turn. *Nature* **318**, 480–483.

82. Ramakrishna, V. & Sasidhar, Y.U. (1997) A pentapeptide model for an early folding step in the refolding of staphylococcal nuclease: The role of its turn propensity. *Biopolymers* **41**, 181–191.

83. Haque, T.S. & Gellman, S.H. (1997) Insights on β-hairpin stability in aqueous solution from peptides with enforced type I' and type II' β-turns. *J. Am. Chem. Soc.* **119**, 2303–2304.

84. de Alba, E., Jimenez, M.A. & Rico, M. (1997) Turn residue sequence determines β-hairpin conformation in designed peptides. *J. Am. Chem. Soc.* **119**, 898–912.

85. Stroup, A.N. & Gierash, L.M. (1990) Reduced tendency to form a β turn in peptides from the P22 tailspike protein correlates with a temperature-sensitive folding defect. *Biochemistry* **29**, 9765–9771.

86. Predki, P.F., Agrawal, V., Brunger, A.T. & Regan, L. (1996) Amino-acid substitutions in a surface turn modulate protein stability. *Nature Struct. Biol.* **3**, 54–58.

87. Munoz, V. & Serrano, L. (1996) Local versus nonlocal interactions in protein folding and stability – an experimentalist's point of view. *Folding & Design* **1**, R71–R77.

88. McGregor, M.J., Islam, S.A. & Sternberg, M.J.E. (1987) Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. *J. Mol. Biol.* **198**, 295–310.

89. Griffiths-Jones, S.R., Sharman, G.J., Maynard, A.J. & Searle, M.S. (1998) Modulation of intrinsic phi,psi propensities of amino acids by neighboring residues in the coil regions of protein structures: NMR analysis and dissection of a beta-hairpin peptide. *J. Mol. Biol.* **284**, 1597–1609.

90. Honig, B. & Cohen, F. (1996) Adding backbone to protein folding: Proteins are polypeptides. *Folding & Design* **1**, R17–R20.

91. Chou, P. & Fasman, G.D. (1978) Empirical predictions of protein conformation. *Annu. Rev. Biochem.* **47**, 251–276.

92. Dill, K.A. & Shortle, D. (1991) Denatured states of proteins. *Annu. Rev. Biochem.* **60**, 795–825.

93. Dyson, H.J. & Wright, P.E. (1993) Peptide conformation and protein folding. *Curr. Opin. Struct. Biol.* **3**, 60–65.

94. Dyson, H.J., Sayre, J.R., Merutka, G., Shin, H.-C., Lerner, R.A. & Wright, P.E. (1992) Folding of peptide fragments comprising the complete sequence of proteins. II. Plastocyanin. *J. Mol. Biol.* **226**, 819–835.

95. Segawa, S.-I., Fukuno, T., Fujiwara, K. & Noda, Y. (1991) Local structures in unfolded lysozyme and correlation with secondary structures in the native conformation: Helix-forming or -breaking propensity of peptide segments. *Biopolymers* **31**, 497–509.

96. Yang, J.J., Buck, M., Pitkeathly, M., Kotik, M., Haynie, D.T., Dobson, C.M. & Radford, S.E. (1995) Conformational properties of four peptides spanning the sequence of hen lysozyme. *J. Mol. Biol.* **252**, 483–491.

97. Sancho, J., Neira, J.L. & Fersht, A.R. (1993) An N-terminal fragment of barnase has residual helical structure similar to that in refolding intermediate. *J. Mol. Biol.* **224**, 749–758.

98. Kemmink, J. & Creighton, T.E. (1993) Local conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor and their roles in folding. *J. Mol. Biol.* **234**, 861–878.

99. Wu, L.C., Laub, P.B., Elove, G.A., Carey, J. & Roder, H. (1993) A non-covalent peptide complex as a model for an early folding intermediate of cytochrome *c*. *Biochemistry* **32**, 10271–10276.

100. Itzhaki, L.S., Neira, J.L., Ruiz-Sanz, J., de Prat-Gay, G. & Fersht, A.R. (1995) Search for nucleation sites in smaller fragments of chymotrypsin inhibitor 2. *J. Mol. Biol.* **254**, 289–304.

101. Blanco, F. & Serrano, L. (1995) Folding of protein GB1 domain studied by conformational characterisation of fragments comprising its secondary structure elements. *Eur. J. Biochem.* **230**, 634–649.

102. Reymond, M.T., Merutka, G., Dyson, H.J. & Wright, P.E. (1997) Folding propensities of peptide fragments of myoglobin. *Protein Sci.* **6**, 706–716.

103. Muñoz, V., Serrano, L., Jiménez, M.A. & Rico, M. (1995) Structural analysis of peptides encompassing all α-helices of three α/β parallel proteins: Che-Y, flavodoxin and P21-ras: Implications for α-helix stability and the folding of α/β parallel proteins. *J. Mol. Biol.* **247**, 648–669.

104. Viguera, A.R., Jimenez, M.A., Rico, M. & Serrano, L. (1996) Conformational analysis of peptides corresponding to beta-hairpins and a beta-sheet that represent the entire sequence of the alpha-spectrin SH3 domain. *J. Mol. Biol.* **255**, 507–521.

105. Alexandrescu, A.T. & Shortle, D. (1994) Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease. *J. Mol. Biol.* **242**, 527–546.

106. Wang, Y. & Shortle, D. (1995) The equilibrium folding pathway of staphylococcal nuclease: Identification of the most stable chain-chain interactions by NMR and CD spectroscopy. *Biochemistry* **34**, 15895–15905.

107. Zhang, O., Kay, L.E., Shortle, D., Forman-Kay, J.D. (1997) Comprehensive NOE characterisation of a partially folded large fragment of staphylococcal nuclease D131D, using NMR *J. Mol. Biol.* **272**, 9–20.

108. Sosnick, T.R., Shtilerman, M.D., Mayne, L. & Englander, S.W. (1997) Ultrafast signals in protein folding and the polypeptide contracted state. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8545–8550.

109. Neira, J.L., Itzhaki, L.S., Ladurner, A.G., Davis, B., de Prat-Gay, G. & Fersht, A.R. (1997) Following co-operative formation of secondary and tertiary structure in a single protein module. *J. Mol. Biol.* **268**, 185–197.

110. Fersht, A.R. (1995) Optimization of rates of protein folding: The nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10869–10873.

111. Evans, P.A., Topping, K.D., Woolfson, D.N. & Dobson, C.M. (1991) Hydrophobic clustering in nonnative states of a protein: Interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. *Proteins* **9**, 248–266.

112. Neri, D., Billeter, M., Wider, G. & Wuthrich, K. (1992) NMR determination of residual structure in urea denatured protein, the 434 repressor. *Science* **257**, 1559–1563.

113. Logan, T.M., Thériault, Y. & Fesik, S.W. (1994) Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J. Mol. Biol.* **236**, 637–648.

114. Arcus, V.L., Vuilleumier, S., Freund, S.M., Bycroft, M. & Fersht, A.R. (1995) A comparison of the pH, urea, and temperature- denatured states of barnase by heteronuclear NMR: Implications for the initiation of protein folding. *J. Mol. Biol.* **254**, 305–321.

115. Pan, H., Barbar, E., Barany, G. & Woodward, C. (1995) Extensive non-random structure in reduced and unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* **34**, 13974–13981.

116. Wong, K.B., Freund, S.M. & Fersht, A.R. (1996) Cold denaturation of barstar: ^{1}H, ^{15}N and ^{13}C NMR assignment and characterisation of residual structure. *J. Mol. Biol.* **259**, 805–818.

117. Schwalbe, H., Fiebig, K.M., Buck, M., Jones, J.A., Grimshaw, S.B., Spencer, A., Glaser, S.J., Smith, L.J. & Dobson, C.M. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry* **36**, 8977–8991.

118. Bond, C.J., Wong, K.-B., Clarke, J., Fersht, A.R. & Dagget, V. (1997) Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: Description of the folding pathway. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13409–13413.

119. Wuthrich, K. (1994) NMR assignments as a basis for structural characterisation of denatured states of globular proteins. *Curr. Opin. Struct. Biol.* **4**, 93–99.

120. Eliezer, D., Yao, J., Dyson, H.J. & Wright, P.E. (1998) Structural and dynamic characterization of partially folded states of myoglobin and implications for protein folding. *Nature Struct. Biol.* **5**, 148–155.

121. Meekhof, A.E. & Freund, S.M. (1999) Probing residual structure and backbone dynamics an the milli- and picosecond timescale in a urea-denatured fibronectin type III domain. *J. Mol. Biol.* **286**, 579–592.

122. Sosnick, T.R., Mayne, L. & Englander, S.W. (1996) Molecular collapse: The rate limiting step in two-state cytochrome *c* folding. *Proteins: Struct. Funct. Genet.* **24**, 413–426.

123. Sosnick, T.R., Jackson, S., Wilk, R., Englander, S.W. & DeGrado, W.F. (1996) The role of helix formation in the folding of a fully α-helical coiled-coil. *Proteins: Struct. Funct. Genet.* **24**, 427–432.

124. Qi, P.X., Sosnick, T.R. & Englander, S.W. (1998) The burst phase in ribonuclease A folding and solvent dependence of the unfolded state. *Nature Struct. Biol.* **5**, 882–884.

125. Roder, H. & Colon, W. (1997) Kinetic role of early intermediates in protein folding. *Curr. Opin. Struct. Biol.* **7**, 15–28.

126. Dadlez, M. & Kim, P.S. (1995) A third native one-disulphide intermediate in the folding of bovine pancreatic trypsin inhibitor. *Nature Struct. Biology* **2**, 674–678.

127. Dadlez, M. & Kim, P.S. (1996) Rapid formation of the native 14–38 disulphide bond in the early stages of BPTI folding. *Biochemistry* **35**, 16153–16164.

128. Dadlez, M. (1997) Hydrophobic interactions accelerate early stages of the folding of BPTI. *Biochemistry* **36**, 2788–2797.

129. Krokoszynska, I., Dadlez, M. & Otlewski, J. (1998) Structure of single-disulphide variants of bovine pancreatic trypsin inhibitor (BPTI) as probed by their binding to bovine beta-trypsin. *J. Mol. Biol.* **275**, 503–513.

130. Creighton, T.E. (1997) Protein folding coupled to disulphide bond formation. *Biol. Chem.* **378**, 731–744.

131. Barbar, E., LiCata, V.J., Barany, G. & Woodward, C. (1997) Local fluctuations and global unfolding of partially folded BPTI detected by NMR. *Biophys. Chem.* **64**, 45–57.

132. Ferrer, M., Barany, G. & Woodward, C. (1995) Partially folded, molten globule and molten coil states of bovine pancreatic trypsin inhibitor. *Nature Struct. Biol.* **2**, 211–217.

133. Amir, D. & Haas, E. (1988) Reduced bovine pancreatic trypsin inhibitor has a compact structure. *Biochemistry* **27**, 8889–8893.

134. Amir, D., Krausz, S. & Haas, E. (1992) Detection of local structures in reduced unfolded bovine pancreatic trypsin inhibitor. *Proteins: Struct. Funct. Genet.* **13**, 162–173.

135. Ittah, V. & Haas, E. (1995) Nonlocal interactions stabilize long range loops in the initial folding intermediates of reduced bovine pancreatic trypsin inhibitor. *Biochemistry* **34**, 4493–4506.

136. Peng, Z.Y., Wu, C.L. & Kim, P.S. (1995) Local structural preferences in the α-lactalbumin molten globule. *Biochemistry* **34**, 3248–3252.

137. Burgering, M.J., Hald, M., Boelens, R., Breg, J.N. & Kaptein, R. (1995) Hydrogen exchange studies of the Arc repressor: Evidence for a monomeric folding intermediate. *Biopolymers* **35**, 217–226.

138. Agashe, V.R., Shastry, M.C.R. & Udgoankar, J.B. (1995) Initial hydrophobic collapse in the folding of barstar. *Nature* **377**, 754–757.

139. Kuwajima, K. (1996) Stopped-flow circular dichroism; in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G.D., ed.) pp. 159–181, Plenum Press, New York.

140. Gronenborn, A.M. & Clore, G.M. (1994) Experimental support for the "hydrophobic zipper" hypothesis. *Science* **263**, 536.

141. Lorch, M., Mason, J.M., Clarke, A.R. & Parker, M.J. (1999) Effects of core mutations on the folding of a β-sheet protein: Implications for backbone organisation in the I-state. *Biochemistry* **38**, 1377–1385.

142. Khorasanizadeh, S., Peters, I.D., Butt, T.R. & Roder, H. (1993) Stability and folding of a tryptophan containing mutant of ubiquitin. *Biochemistry* **32**, 7054–7063.

143. Khorasanizadeh, S., Peters, I.D. & Roder, H. (1996) Evidence for a three state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nature Struct. Biol.* **3**, 193–205.

144. Colon, W., Elove, G.A., Wakem, L.P., Sherman, F. & Roder, H. (1996) Side-chain packing of the N-terminal and C-terminal helices play a critical role in the kinetics of cytochrome *c* folding. *Biochemistry* **35**, 5538–5549.

145. Gast, K., Chaffotte, A.F., Zirwer, D., Guillou, Y., Mueller-Frohne, M., Cadieux, C., Hodges, M., Damaschun, G. & Goldberg, M.E. (1997) Lack of coupling between secondary structure formation and collapse in a model polypeptide that mimics early folding intermediates, the F_2 fragment of the *Escherichia coli* tryptophan-synthase β chain. *Protein Sci.* **6**, 2578–2588.

146. Hamada, D., Segawa, S. & Goto, Y. (1996) Non-native alpha-helical intermediate in the refolding of beta-lactoglobulin, a predominantly beta-sheet protein. *Nature Struct. Biol.* **3**, 868–873.

147. Blanco, F.J., Serrano, L. & Forman-Kay, J.D. (1998) High population of non-native structures in the denatured state are compatible with the formation of the native folded state. *J. Mol. Biol.* **284**, 1153–1164.

148. Prieto, J., Wilmans, M., Jimenez, M.A., Rico, M. & Serrano, L. (1997) Non-native local interactions in protein folding and stability: Introducing a helical tendency in the all β-sheet α-spectrin SH3 domain. *J. Mol. Biol.* **268**, 760–778.

149. Minor, D.L. & Kim, P.S. (1996) Context-dependent secondary structure formation of a designed sequence. *Nature* **380**, 730–734.

150. Anfinsen, C.B. (1972) The formation and stabilisation of protein structure. *Biochem. J.* **128**, 737–749.

151. Wright, P.E., Dyson, H.J. & Lerner, R.A. (1988) Conformation of peptide fragments of proteins in aqueous solutions: Implications for initiation of protein folding. *Biochemistry* **27**, 7167–7175.

152. Freund, S.M., Wong, K.B., Fersht, A.R. (1996) Initiation sites of protein folding by NMR analysis. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10600–10603.

153. Ptitsyn, O.B. (1981) Protein folding: General physical model. *FEBS Lett.* **131**, 197–201.

154. Kaiser, E.T. & Kezdy, F.J. (1984) Amphiphilic secondary structure; Design of peptide hormones. *Science* **223**, 249–255.

155. Kim, P.S. & Baldwin, R.L. (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**, 459–489.

156. Kim, P.S. & Baldwin, R.L. (1990) Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631–660.

157. Harrison, S.C. & Durbin, R. (1985) Is there a single pathway for the folding of a

polypeptide chain? *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4028–4030.

158. Karplus, M. & Weaver, D.L. (1994) Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Sci.* **3**, 650–668.

159. Wetlaufer, D.B. (1973) Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **70**, 697–701.

160. Ptitsyn, O.B. (1993) Stage mechanism of the self-organisation of protein molecules. *Dokl. Akad. Nauk* **210**, 1213–1215.

161. Plaxco, K.W., Simons, K.T. & Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994.

162. Creighton, T.E. (1988) On the relevance of non-random polypeptide conformations for protein folding. *Biophys. Chem.* **31**, 155–162.

163. Creighton, T.E. (1983) An empirical approach to protein conformation stability and flexibility. *Biopolymers* **22**, 49–58.

164. Baldwin, R.L. (1996) Why is protein folding so fast? *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2627–2628.

165. Abkevich, V.I., Gutin, A.M., Shakhnovich, E.I. (1996) How the first biopolymers could have evolved. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 839–844.

166. Govindarajan, S. & Goldstein, R.A. (1995) Optimal local propensities for model proteins. *Proteins* **22**, 413–418.

167. Doyle, R., Simons, K., Qian, H. & Baker, D. (1997) Local interactions and the optimisation of protein folding. *Proteins* **29**, 282–291.

168. Unger, R. & Moult, J. (1996) Local interactions dominate folding in a simple protein model. *J. Mol. Biol.* **259**, 988–994.

169. Stickle, D.F., Presta, L.G., Dill, K.A. & Rose, G.D. (1992) Hydrogen bonding in globular proteins. *J. Mol. Biol.* **226**, 1143–1159.

170. Plaxco, K.W. & Baker, D. (1998) Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13591–13596.

171. Blanco, F.J., Angrand, I. & Serrano, L. (1999) Exploring the conformational properties of the sequence space between two proteins with different folds: An experimental study. *J. Mol. Biol.* **285**, 741–753.

172. Cordes, M.H.J., Walsh, N.P., McKnight, J. & Sauer, R. (1999) Evolution of a protein fold *in vitro*. *Science* **284**, 325–327.

173. Brown, B.M. & Sauer, R.T. (1999) Tolerance of Arc repressor to multiple-alanine substitutions. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1983–1988.

174. Xiong, H., Buckwalter, B.L., Shieh, H.-M. & Hecht, M.H. (1995) Periodicity of polar and nonpolar amino acids is the major determinant of secondary structure in self-assembling oligomeric peptides. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6349–6353.

175. Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M. & Hecht, M.H. (1993) Protein design by binary patterning of polar and nonpolar amino acids. *Science* **262**, 1680–1685.

176. O'Shea, E.K., Klemm, J.D., Kim, P.S. & Alber, T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled-coil. *Science* **254**, 539–544.

177. Munoz, V., Cronet, P., Lopez-Hernandez, E. & Serrano, L. (1996) Analysis of the effect of local interactions in protein stability. *Folding & Design* **1**, 167–178.

178. Cavangero, S., Dyson, H.J. & Wright, P.E. (1999) Effect of H helix destabilizing mutations on the kinetic and equilibrium folding of apomyoglobin. *J. Mol. Biol.* **285**, 269–282.

179. Shin, H.-C., Merutka, G., Waltho, J.P., Tennant, L.L., Dyson, H.J. & Wright, P. (1993) Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin. *Biochemistry* **32**, 6356–6364.

180. Viguera, A.R., Blanco, F.J. & Serrano, L. (1995) The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. *J. Mol. Biol.* **247**, 670–681.

181. Villegas, V., Viguera, A.R., Aviles, F.X. & Serrano, L. (1996) Stabilisation of proteins by rational design of β-helix stability using helix/coil transition theory. *Folding & Design* **1**, 29-34.

182. Viguera, A.R., Villegas, V., Aviles, F.X. & Serrano, L. (1997) Favorable native-like helical local interactions can accelerate protein folding. *Folding & Design* **2**, 23-33.

183. Ladurner, A.G., Itzhaki, L.S., Dagget, V. & Fersht, A.R. (1998) Synergy between simulation and experiment in describing the energy landscape of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8473-8478.

184. Chiti, F., Taddei, N., Webster, P., Hamada, D., Fiaschi, T., Ramponi, G. & Dobson, C. (1999) Acceleration of the folding of acylphosphatase by stabilisation of local secondary structure. *Nature Struct. Biol.* **6**, 380-387.

185. Viguera, A.-R. & Serrano, L. (1997) Loop length, intramolecular diffusion and protein folding. *Nature Struct. Biol.* **4**, 939-946.

186. Neira, J.L. & Fersht, A. (1999) Exploring the folding funnel of a polypeptide chain by biophysical studies on protein fragments. *J. Mol. Biol.* **285**, 1309-1333.

187. Brunet, A.P., Huang, E.S, Huffine, M., Loeb, J.E., Weltman, R.J. & Hecht, M.H. (1993) The role of turns in the structure of an alpha-helical protein. *Nature* **364**, 355-358.

188. Zhou, H.X., Hoess, R.H. & DeGrado, W.F. (1996) *In vitro* evolution of thermodynamically stable turns. *Nature Struct. Biol.* **3**, 446-451.

189. Castagnoli, L., Vetriani, C. & Cesareni, C. (1994) Linking an easily detectable phenotype to the folding of a common structural motif. Selection of a rare mutations that prevent the folding of Rop. *J. Mol. Biol.* **234**, 378-387.

190. Ybe, J.A. & Hecht, M.H. (1996) Sequence replacements in the central β-turn of plastocyanin. *Protein Sci.* **5**, 814-824.

191. Villegas, V., Martinez, J.C., Aviles, F.X. & Serrano, L. (1998) Structure of the transition state in the folding process of human procarboxydase A2 activation domain. *J. Mol. Biol.* **283**, 1027-1936.

192. Goldberg, J.M. & Baldwin, R.L. (1999) A specific transition state for S-peptide combining with folded S-protein and then refolding. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2019-2024.

193. Zwanzig, R. (1997) Two-state models for protein folding kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 148-150.