Resolving local structures and dynamics of proteins with 2D IR spectroscopy
Roy, Santanu

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2012

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 6

Temperature-jump relaxation kinetics of a $\beta$-hairpin

We introduce a simulation protocol to model a temperature jump (T-jump) experiment combined with linear infrared absorption and two-dimensional infrared spectroscopy. For complex protein molecules it is often hard to interpret the infrared spectral changes observed in such experiments. Simulating T-jump experiments which allow to analyze the infrared spectra of proteins subject to the T-jump, is needed to understand the relaxation kinetics and the underlying physical process. In this chapter, we propose a new strategy for that and apply it to a cyclic $\beta$-hairpin peptide. The spectral changes in the relaxation process after the T-jump are probed and a sub-nanosecond kinetics of the peptide side chain is observed. Determining the unfolding kinetics is not possible as it turns out that the fully unfolded configuration is not accessible with the force field employed here.
Chapter 6: Temperature-jump relaxation kinetics of a β-hairpin

6.1 Introduction

Research activities on the protein folding problem [8] have made great progress in last decades with the advances of experimental [21, 116, 119–121] and computational techniques [26–28]. This problem originated from a thermodynamic hypothesis [10, 11] which states that given an amino acid sequence a protein folds into a thermodynamically stable unique three-dimensional native structure. "What is the pathway followed in a folding process to reach the native structure?" is one of the major questions addressed in this context. In the modern view, protein folding is described as an ensemble of many parallel pathways going through several transitions among different intermediate structural sub-ensembles [9]. This leads to the question: "What are the transition rates among these sub-ensembles required to describe the folding/unfolding kinetics?" Simple statistical mechanical models are used to determine the kinetics of peptides studied with fluorescence [194, 195] and linear infrared absorption [72], and nuclear magnetic resonance (NMR) [17, 18, 207, 208] spectroscopies. Recent developments on single molecule spectroscopic techniques [22, 34, 209] have allowed to track the folding pathways and determine the corresponding kinetics at the single molecule level. To distinguish between different folding intermediates with unique hydrogen bonding structures, two-dimensional infrared (2D IR) spectroscopy [35] is an ideal tool to apply. This technique has been applied to study protein misfolding and aggregation such as the amyloid fibril formation [74] and to the temperature-dependent study of protein folding [42, 50, 73, 133, 210]. On the other hand, as a computational tool, molecular dynamics simulation is used for theoretical investigation of structure, dynamics, and kinetics of proteins and has been applied to study folding of fast-folding proteins [28]. The recently developed Markov state model [58, 59, 211] is a useful tool to determine the kinetics between the metastable structures accessed in MD simulations.

The infrared experimental spectra of proteins are often hard to interpret, because their complex structures and dynamics result in complex congested spectral shapes. Thus, theoretical models are required to identify the underlying physical phenomena in experimental observations. Here, we introduce a simulation protocol to model a temperature-jump (T-jump) experiment combined with linear infrared absorption and 2D IR spectroscopy, which can be used to unravel the unfolding kinetics of proteins and peptides. In a T-jump experiment for a protein a sudden thermal change is brought into its surroundings, followed by the relaxation of the protein to
the new thermal equilibrium. Experimental T-jump studies in combination with linear infrared absorption and 2D IR spectrosopies have revealed the folding/unfolding rates for a number of proteins and peptides [42, 50, 70–73, 133, 210, 212, 213]. To bridge the gap between experiment and theory, a T-jump simulation methodology using Markov states was reported earlier [214]. This devised a protocol to simulate a T-jump experiment, but did not provide a way to give a detailed physical interpretation of the underlying kinetics determining the spectral changes in relaxation process. How the kinetics between the Markov states is related to the relaxation rates measured in a T-jump experiment, was not determined. To fill this void, here, we introduce a new T-jump simulation protocol. This protocol is based on a thorough conformational search using replica-exchange molecular dynamics (REMD) simulations [56, 57, 215, 216] to obtain structural sub-ensembles. These sub-ensembles are used for modeling a T-jump experiment. We establish a connection of the kinetics between these sub-ensembles to the spectral changes observed in the T-jump simulation, explaining the underlying physical phenomena. We will investigate whether this kinetics is related to the folding/unfolding rates measured in a real T-jump experiment.

To understand protein folding it is crucial to have deep insights into the three-dimensional structure of proteins. The secondary structural motifs of proteins, the α-helix and the β-sheet, are generally considered as the units for developing models to investigate protein folding. Here, to build up the simulation methodology for a T-jump experiment, we consider a cyclic β-hairpin peptide which is an analog of gramicidin S. Gramicidin S is a very stable cyclic decapptide with anti-parallel β-sheet structure bordered with two type II' β-turns as revealed by X-ray crystallography [217], NMR spectroscopy [218], gas phase linear infrared absorption [127, 219], and 2D IR spectroscopy [43]. Water soluble analogs of this antibiotic peptide were designed for circular dichroism and NMR spectroscopic investigation of the effects of amino acid substitution on the beta-sheet structure and the turn [31]. Like for the original gramicidin, the turn is type II’ for the analog with 10 amino acids, GS10 (cyclo(VKLYPVKLYP)) (see Fig. 6.1). A linear infrared absorption experiment combined with a T-jump disclosed that folding/unfolding time scales for this analog are around 140 nanoseconds (ns) [72].

Because of the fast kinetics and the rigidity of the very stable peptide, GS10, it should be possible to explore most of its configuration space and determine the kinetics among the structural sub-ensembles with MD sim-
Chapter 6: Temperature-jump relaxation kinetics of a $\beta$-hairpin

Figure 6.1: The backbone structure of GS10: the red circles indicate the amide units preceding proline (Y4 and Y9), which are called the tertiary amide units. The other units are called the secondary amide units. Because of the ring in proline, a tertiary amide unit is structurally distinct from a secondary amide unit.

ulations. Thus, we expect that GS10 is an ideal system to investigate a new methodology for modeling a T-jump experiment. Advanced clustering techniques combined with Markov states modeling [58] used in the earlier proposal for the T-jump simulation [214] are not needed for such a small, rigid, and stable cyclic peptide, because the corresponding configuration space is expected to have a small size. Therefore, here, we probe the backbone structure involved in cross-strand H-bonding and side chain structures for this peptide as determined from an extensive REMD simulation. Theoretical investigations combined with linear infrared absorption spectroscopy have been carried out on this peptide [220, 221], but the kinetics between different structural components have not been reported and a T-jump experiment accompanied by 2D IR spectroscopy has been neither performed nor simulated for this peptide. In this chapter, we theoretically examine to what extent the linear infrared absorption and 2D IR spectra are sensitive to conformational changes of the peptide subject to a T-jump, and determine the underlying kinetics.

GS10 contains ten amide I oscillators, each one is associated with predominantly a CO stretch vibration. The amide units of tyrosine (Y4/Y9) preceding proline (P5/P10) are tertiary amide units (Fig. 6.1), the other amide I units are secondary ones. The amide unit preceding proline is an exceptional amide I unit of which the gas phase vibrational frequency is 27 cm$^{-1}$ lower than the secondary amide I oscillators [142]. Thus, the peak originating from an amide I unit preceding proline is expected to be located in the red side of absorption spectra of peptides containing proline [98]. Based on a newly developed parametrization for the amide I vibra-
tional frequencies in proteins containing proline [98], a recent combined study of MD simulation and 2D IR spectroscopy has shown that by probing the amide I vibration of the unit preceding proline one can distinguish between different turn structures induced by mutation [222]. Formation of turns in proteins are crucial steps in protein folding and identifying different types of turn configurations in a folding process with 2D IR spectroscopy combined with MD simulations allows to understand how the turns form.

In this chapter, we choose to probe the vibrational dynamics of the amide units preceding proline located in the turns by simulating linear and 2D IR spectra of different structural sub-ensembles of GS10. We examine how the spectral region corresponding to the amide units preceding proline change in the relaxation process after the peptide experiences a T-jump.

The remainder of this chapter is organized as follows. We describe the theory used in modeling a T-jump in Section 6.2. The details of MD and spectral simulations are given in Section 6.3. Discussions of results and concluding remarks are presented in Section 6.4 and Section 6.5, respectively.

### 6.2 Simulation protocol

In a T-jump experiment the water solution of a peptide or protein is heated by exciting the overtone transition of the O-H (or O-D) stretch vibrations of water using a an intense near infrared laser pulse [72, 212]. The laser pulse duration for heating is typically < 10 ns. If the kinetics of a protein is faster than the time period for heating, the corresponding dynamics cannot be resolved by a T-jump experiment. To give a simple illustration of what happens if a peptide with two structural sub-ensembles ($S_1$ and $S_2$) is subject to a T-jump, the one dimensional free energy landscape of the system as a function of a reaction coordinate is depicted in Fig. 6.2. A reaction coordinate describes the path to monitor the transition between $S_1$ and $S_2$. The relative free energy ($A$) in a canonical ensemble at temperature $T$, is related to the occurrence of $S_1$ ($\Omega_{S_1}$) and $S_2$ ($\Omega_{S_2}$), i.e. $A = -k_B T \ln(\Omega_{S_1}/\Omega_{S_2})$, where, $k_B$ is the Boltzman constant. The shape of the free energy landscape changes if the temperature is changed, because at a particular temperature each state has a particular equilibrium population. Let us consider that at an initial temperature, $T_i$, $S_1$ has more population than $S_2$. Now, we make the assumption that no structural changes of the peptide occur during the jump to a new temperature, $T_f$. 


Figure 6.2:  *Left panel:*  Schematics representation of the one-dimensional free energy landscape of a system with two states, $S_1$ and $S_2$, at temperature $T_i$. The blue lines filling the minima are used to indicate the equilibrium relative populations of $S_1$ and $S_2$. *Right panel:*  Assuming the populations do not change during the T-jump from $T_i$ to $T_f$, these evolve with time following the T-jump to relax to the new free energy landscape in thermal equilibrium at the final temperature, $T_f$.

that is, the population of $S_1$ and $S_2$ do not change. After the T-jump, the system will relax to the new thermal equilibrium corresponding the new free energy landscape, where $S_2$ is more populated than $S_1$. Below, we present a model to simulate a T-jump experiment for a system with $N$ structural sub-ensembles that constitute its configuration space.

The configuration space of a system (peptide or protein) can be generated by performing MD simulations of the system. In a constant temperature MD simulation it is possible that all possible configurations may not be accessed, the MD trajectory may get stuck in kinetics traps on the free energy surface. REMD is a very useful simulation technique to overcome this problem. Using this technique multiple MD simulations of a protein run in parallel at exponentially distributed temperatures and replicas of the protein at two adjacent temperatures are exchanged with a certain probability based on the standard Metropolis criteria [57]. For example, if $T_i$ and $T_j$ are the temperatures of the two adjacent replicas $i$ and $j$ and $U_i$ and $U_j$ are their corresponding potential energies, then the probability that these replicas will exchange is:

$$P(T_i \leftrightarrow T_j) = \min(1, e^{\Delta})$$

(6.1)

Where,

$$\Delta = (\beta_i - \beta_j)(U_i - U_j)$$

(6.2)
Here, $\beta_i = k_B T_i$ and $\beta_j = k_B T_j$. The main purpose of REMD is to make configurations at high temperatures available to the MD simulations at low temperatures and vice versa, allowing to access most part of the configuration space by passing over the activation barriers at higher temperatures.

The configuration space can be divided into a number ($N_s$) of structural sub-ensembles (states) that are defined by local minima on the potential surface of the system. When the system is in a canonical ensemble at temperature $T_i$, the populations of the states in this thermal equilibrium are $P_{T_i}^1$, $P_{T_i}^2$, ......., $P_{T_i}^{N_s}$, where, $\sum_{j=1}^{N_s} P_{T_i}^j = 1$, and $0 \leq P_{T_i}^j \leq 1$. The states interconvert between each other with specific rates that determine the equilibrium at a specific temperature. We assume that the system subject to a T-jump ($\Delta T$) in short period of time ($\delta t$) does not undergo structural changes during that period. After the T-jump, the system will follow a relaxation process to get to a new equilibrium at the final temperature, $T_f = T_i + \Delta T$. We assume that the populations of the states remain unchanged during the T-jump and evolve afterwards with time depending on the kinetics at the final temperature, $T_f$.

The master equation for the kinetics at the temperature $T_f$ is

$$\frac{dP_{T_f}^j}{dt} = -P_{T_f}^j \sum_{k \neq j}^{N_s} \lambda_{jk} + \sum_{k \neq j}^{N_s} P_{T_f}^k \lambda_{kj}. \quad (6.3)$$

$P_{T_f}^j$ is the population of state $j$, and $\lambda_{jk}$ is for the rate at which state $j$ converts to state $k$. This can also be formulated as a matrix differential equation:

$$\frac{dP_{T_f}}{dt} = \Gamma P_{T_f} \quad (6.4)$$

Where, $P = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_{N_s} \end{pmatrix}$ and $\Gamma = \begin{pmatrix} -\sum_{k \neq 1}^{N_s} \lambda_{1k} & \lambda_{12} & \cdots & \lambda_{1N_s} \\ \lambda_{12} & -\sum_{k \neq 2}^{N_s} \lambda_{2k} & \cdots & \lambda_{2N_s} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N_s2} & \cdots & \cdots & -\sum_{k \neq N_s} \lambda_{N_sN_s} \end{pmatrix}$

The solution of Eq. 6.4 is

$$P_{T_f}(t) = \exp(\Gamma t)P_{T_f}(0). \quad (6.5)$$
The initial population vector, \( P_{T_{f}}(0) \), is the same as the population vector before the T-jump, \( P_{T_{i}} \), because of the assumption that the populations of sub-ensembles do not change during the T-jump.

Linear infrared absorption and 2D IR spectra of the system can then be obtained as the sum of the spectra of all states weighted according to their populations. If \( I_{T_{f}}^{j} \) is the spectrum of state \( j \) at the final temperature \( T_{f} \), the weighted spectrum at this temperature for the system is

\[
I_{T_{f}}^{T_{f}}(t) = \sum_{j} P_{T_{f}}^{T_{f}}(t) I_{T_{f}}^{T_{f}}. \tag{6.6}
\]

In the relaxation process following the T-jump from \( T_{i} \) to \( T_{f} \) the spectral change is expressed as \( \Delta I \) given by

\[
\Delta I(t) = I_{T_{f}}^{T_{f}}(t) - \sum_{j} P_{T_{i}}^{T_{i}} I_{T_{i}}^{T_{i}}. \tag{6.7}
\]

Here, \( I_{T_{i}}^{T_{i}} \) is the spectrum of state \( j \) when the system is in the thermal equilibrium at the initial temperature, \( T_{i} \). In Fig. 6.3, we pictorially demonstrate the simulation protocol for a T-jump experiment.

### 6.3 Methods

#### 6.3.1 REMD and the Configuration Space

The NMR structure [31] of GS10 was chosen as the initial structure for the REMD simulation and it was solvated in D\(_2\)O. All MD simulation reported later in this study were also performed in D\(_2\)O. This is to be consistent with infrared experiments, where D\(_2\)O is used instead of H\(_2\)O to avoid the overlapping of water bending vibrations in H\(_2\)O with the amide I vibrations in proteins. The acidic protons of GS10 were replaced with deuterium. The formal charge of GS10 is +2 because it has two lysine amino acid residues with positively charged side chains. Therefore, two Cl\(^-\) ions were added in the simulation to keep the system neutral in simulation. 32 replicas of the system were obtained at 32 temperatures generated between 273 K and 410 K with the exchange probability of 0.2 between two neighboring replicas [223]. The time step used for the 32 parallel MD simulations at constant volume was 2 fs, and attempts were made every 1 ps for the exchange
The simulation protocol for a T-jump experiment: First, REMD simulations are performed to explore the configuration space of a peptide. Analyzing the REMD trajectories structurally distinct sub-ensembles, $S_1, \cdots, S_n$ are obtained. Now, to model a T-jump from temperature $T_i$ to temperature $T_f$, the kinetics, i.e. the transition rates among the sub-ensembles at $T_f$ are determined. On the other hand, the spectrum of each sub-ensemble at both temperatures is calculated. The weighted spectrum at $T_i$ is obtained accounting for the REMD populations of the sub-ensembles at that temperature. Based on the kinetics, the population of each sub-ensemble is determined as a function of time following the T-jump. The spectra calculated at $T_f$ are weighted according to the time-dependent populations, and then the weighted spectrum as a function of time is obtained. We note that the populations immediately after the T-jump are assumed to be the same as those before the T-jump. Finally, the difference between the weighted spectra at different time points after the T-jump and the weighted spectrum at $T_i$ are calculated.
between two adjacent replicas. The all-atom OPLS [150] and SPC/E force fields [151] were used to describe the peptides and the water, respectively. The LINCS algorithm [152] was employed to constrain all bond lengths. The reaction field method based long range electrostatic interactions were treated within a cut-off of 1.4 nm. The Nosé-Hoover thermostat [153, 154] was used to couple the replicas to the constant temperature baths.

Each replica was simulated for 468 ns and the last 456 ns of every trajectory was used for the production run. Thus, a total length of 14.6 µs was used to explore the configuration space. All simulations were performed with Gromacs-4.0.7 [149].

Analyzing the REMD trajectories we found that we could categorize the structural sub-ensembles for this peptide in two ways: (a) the first is based on the distribution of distances between cross-strand amino acids describing the backbone and H-bonding structures, (b) the second is related to side
chain orientation of the two lysine amino acid residues. In Fig. 6.4a four possible cross-strand H-bonds are depicted. To identify H-bonding states we introduced four H-bonding distance parameters: $d_1$, $d_2$, $d_3$, and $d_4$. From the distribution of these distance parameters (Fig. 6.4c) at 330 K we see that the H-bonds in the turn continuously form and break, whereas the middle strand hydrogen bonds remain intact for this cyclic peptide. This is true for the complete temperature range that we have investigated. Thus, here, we constitute four H-bonding states corresponding to four minima in the two-dimensional relative free energy surface (at 330 K) as a function of only $d_1$ and $d_4$ ignoring $d_2$ and $d_3$ as presented in Fig. 6.4d: $S_{1}^{HB}$, $S_{2}^{HB}$, $S_{3}^{HB}$, and $S_{4}^{HB}$. The $S_{3}^{HB}$ state is the global minimum (0 kJ/mol) on the free energy surface. Except for $S_{2}^{HB}$ the energy barrier between these states is within 5 kJ/mol. $S_{2}^{HB}$ is least likely to occur because of its high relative free energy. By rotational symmetry $S_{1}^{HB}$ and $S_{4}^{HB}$ are the same states and together we denote them the intermediate H-bonding state, $I^{HB}$. Representative structures of these states are depicted in Fig. 6.5.

To identify the states corresponding to the side chain orientation of lysine, we calculated the angle between the side chain nitrogen atom (NZ), the backbone $C_{\alpha}$ atom, and the backbone nitrogen atom (N) of lysine (Fig. 6.4b). Since there are two lysine amino acid residues, we have two angles ($\theta_1$ and $\theta_2$). In the two-dimensional relative free energy surface obtained at 330 K as a function of these angles (Fig. 6.4e) we can clearly distinguish four states: $S_{1}^{\theta}$, $S_{2}^{\theta}$, $S_{3}^{\theta}$, and $S_{4}^{\theta}$. Here, $S_{2}^{\theta}$ is the global minimum (0 kJ/mol) on the free energy surface. The barriers between these states are in the range of 5 to 10 kJ/mol. $S_{1}^{\theta}$ and $S_{4}^{\theta}$ are identical by rotational symmetry and therefore, these states together constitute an intermediate state, $I^{\theta}$. We denote these states as the $\theta$-states and representative structures are depicted in Fig. 6.5.

We note here that the H-bonding states and the $\theta$-states are independent of each other. As shown in the free energy landscape for these parameters (Fig. 6.6), we observe hardly any correlation between the H-bonding distances and the orientation angle of the lysine side chains.

Every state has a certain occurrence at a particular temperature. The population of all the states are shown in Fig. 6.7 as a function of temperature. We find that the population of state $S_{3}^{HB}$ which has all H-bonds intact decreases with increasing temperature. The population of the states with broken H-bonds in the turn ($S_{1}^{HB}$, $S_{2}^{HB}$ and $S_{4}^{HB}$) increases as temperature increases. The decay or growth of the populations are monotonic and quantitatively small as a function of temperature which could be at-
Figure 6.5: The H-bonding states and the states with different orientation of the lysine side chains (θ-states).

distributed to the rigidity and stability of the cyclic peptide. For the θ-states, the populations vary as depicted in Fig. 6.7b. The population of $S^\theta_1$ and $S^\theta_4$ remain almost unchanged at all temperatures, whereas the population of $S^\theta_3$ increases and the population of $S^\theta_2$ decreases with increasing temperature. It is important to point out that the populations of $S^\text{HB}_1$ (or $S^\theta_1$) are almost identical to those of $S^\text{HB}_4$ (or $S^\theta_4$) for all temperatures, this is expected by symmetry. Therefore, we conclude that the REMD simulations have converged and the configuration space has been sufficiently sampled. However, one cannot exclude the fact that other stable structures exist, which are kinetically not accessible from our initial structure.
6.3: Methods

Figure 6.6: **Left panel:** Two-dimensional relative free energy landscape (kJ/mol) for one of the parameters defining the H-bonding states \(d_1\) and one of the parameters defining the \(\theta\)-states \(\theta_1\) at 330 K. The contours are spaced by 0.8 kJ/mol starting with 0 kJ/mol. **Right panel:** Slices through the surface for \(\theta_1 = 99^\circ\) (black), \(\theta_1 = 115^\circ\) (red), and \(\theta_1 = 132^\circ\) (green). This shows that the dependency of the free energy on \(d_1\) is hardly affected if \(\theta_1\) is changed, except for the almost systematic shifts of the slice. The observation is the same for other parameters defining the H-bonding states and the \(\theta\)-states, therefore, the H-bonding states are independent of the \(\theta\)-states.

6.3.2 Kinetics

To obtain transition rates between the states at 330 K, five MD simulations in NVT ensembles are performed, where the initial configurations are chosen arbitrarily from the REMD simulations. Each MD trajectory has a length of 390 ns and the last 378 ns of every trajectory is used for the production run, resulting in a total of 1.89 \(\mu\)s \(5 \times 378\) ns) used for analysis. To be consistent, the force fields and the methods used in these MD simulations are kept the same as those for the REMD simulation and are also used for further simulations reported later in this chapter.

In the MD trajectories we find that the system visits different states and the time intervals for multiple transitions between two states are exponentially distributed. For example, if there are \(N\) transitions between state \(S_i\) and \(S_j\), the histogram of the \(N\) time intervals gives an exponentially decaying function. From this function we find the time scale \(\tau_{ij}\) at which transitions occur between the two states. Then the transition rate is given by \(\lambda_{ij} = \tau_{ij}^{-1}\). In the appendix, the histograms corresponding to the transitions between the \(\theta\)-states are depicted.
Chapter 6: Temperature-jump relaxation kinetics of a β-hairpin

Figure 6.7: Temperature dependence of populations determined from the REMD simulation for all the states.

Figure 6.8: Sub-nanosecond kinetics of the θ-states at 330 K (values are in ps) (a). Time evolution of the populations of θ-states in the relaxation process after a T-jump from 300 K to 330 K (b).

From the two-dimensional free energy surface of the H-bonding states (Fig. 6.4d) at 330 K we find that transitions occur mainly between \( S^\text{HB}_3 \) and the intermediate state, \( I^\text{HB} \). The time scale for the transition from \( S^\text{HB}_3 \) to \( I^\text{HB} \) is 80 ps, whereas, it is 20 ps for the reverse transition. For the θ-states the kinetics at 330 K is pictorially presented in Fig. 6.8a. The transitions from \( I^\theta \) to \( S^\theta_2 \) (or \( S^\theta_3 \)) to are more frequent than the reverse transitions. The least frequent transitions are from \( S^\theta_3 \) to \( I^\theta \) (\( \tau \approx 300 \) ps). By symmetry, the time scales of transitions between \( S^\theta_1 \) and other states must be identical as those between \( S^\theta_4 \) and other states. Due to sampling limitations this is not exactly fulfilled; the maximum deviation is \( \sim 15\% \), which is observed when these states convert to \( S^\theta_2 \). In these cases, the average values are considered for further analysis.

We choose the θ-states for simulating a T-jump from 300 K to 330 K, because the time scales for the H-bonding states are so short (maximum is
80 ps) that we can assume that these states are equilibrated within the lifetime of the \( \theta \)-states (minimum time scale is 110 ps). The kinetics between the \( \theta \)-states may not be resolved in a real T-jump experiment as it is of the order of sub-nanoseconds, smaller than the typical time duration for a T-jump. Nevertheless, in this study, the main purpose is to establish the relation between the kinetics and the spectral changes in the relaxation process after the T-jump and the consideration of the \( \theta \)-states is expected to fulfill the purpose.

The initial population vector \((P_T^T(0) = P_T^i, \text{see Section 6.2})\) for the \( \theta \)-states is taken at 300 K determined from the REMD populations. Depending on the transition rates among the \( \theta \)-states at 330 K the populations evolve with time constants as shown in Fig. 6.8b. The population of the intermediate state, \( I^\theta \), is almost invariant (0.9 % increase) in the relaxation process, which is consistent with the REMD simulation where the population of \( S_1^\theta \) and \( S_4^\theta \) hardly change with temperature. The population change from the initial value to the equilibrium one is -3\% for \( S_2^\theta \) and is 2\% for \( S_3^\theta \). We later investigate whether these small changes are sufficient to display spectral changes.

### 6.3.3 Spectral Modeling

The spectral calculations followed a series of steps. First, MD simulations of \( S_1^\theta \), \( S_2^\theta \), \( S_3^\theta \), and \( S_4^\theta \) were performed at 300 K and 330 K. For the MD simulation of a particular \( \theta \)-state, five initial configurations were chosen from the central region of the local minimum corresponding to that \( \theta \)-state on the two-dimensional free energy surface. Then, five trajectories, each with 100 ps length, were obtained resulting in a total length of 500 ps. Snapshots were saved every 20 fs. These trajectories were then used to construct the following floating oscillator Hamiltonian [25].

\[
H(t) = \sum_{i=1}^{N} \left[ \omega_i(t) b_i^\dagger b_i - \frac{\Delta_i(t)}{2} b_i^\dagger b_i^\dagger b_i b_i \right] + \sum_{i,j}^{N} J_{ij}(t) b_i^\dagger b_j + \sum_{i=1}^{N} \vec{\mu}_i(t) \cdot \vec{E}(t) \left[ b_i^\dagger + b_i \right]
\]  

(6.8)

Here, the amide I oscillators are labeled with \( i \). The Bosonic creation and annihilation operators for the amide I vibration on site \( i \) are \( b_i^\dagger \) and
Chapter 6: Temperature-jump relaxation kinetics of a β-hairpin

\( b_i \), respectively. \( \omega_i(t) \) and \( \Delta_i(t) \) are the fluctuating frequency and anharmonicity, respectively, of site \( i \). \( J_{ij}(t) \) is the fluctuating coupling between the \( i^{th} \) and the \( j^{th} \) site. The external electric field \( \vec{E}(t) \) interacts with the system through its transition dipole \( \vec{\mu}_i(t) \).

The frequencies (\( \omega_i \)) for the secondary amide units were found with a Stark effect based approach using the electric field and its gradient on the \( C, O, N \) and \( D \) atoms \[86\]. For the amide I unit preceding proline the field components were considered on the \( C, O, N \) and \( C_\delta \) atoms \[98\]. A Ramachandran angle based nearest-neighbor frequency shift map \[87, 98\] was used to correct these frequencies accounting for through bond effects from nearest neighbor units. For the nearest-neighbor couplings the Ramachandran angle based scheme proposed in Refs. \[87, 98\] was used. The couplings between non-nearest neighbor units were found using the transition charge coupling scheme \[87, 98\]. Finally, the anharmonicity was set to the experimentally determined value, \( 16 \, \text{cm}^{-1} \), for all units \[25\] to describe doubly excited states. A systematic \( 14 \, \text{cm}^{-1} \) red-shift of frequencies for all amide I sites except the one preceding proline was applied, because the frequency maps for the secondary amide unit overestimate its frequency by the same amount \[98\].

The time dependent Hamiltonian (Eq. 6.8) was used for spectral calculations with the Numerical Integration of Schrödinger Equation scheme (NISE) \[100–102\]. In this scheme the time independent Schrödinger equation is solved for short time intervals (10 fs) during which the Hamiltonian is considered to be constant. The time evolution for longer time is determined by the successive propagation in the short time intervals. The linear and nonlinear response functions (rephasing \( k_I \) + nonrephasing \( k_{II} \)) are then calculated to obtain linear infrared absorption and 2DIR spectra, respectively. An \textit{ad hoc} vibrational lifetime (1 ps) is accounted for \[105\] in the spectra calculations. For the 2DIR spectra we used a waiting time of 0 fs. This is the time between the pump and the probe at frequencies \( \omega_1 \) and \( \omega_3 \), respectively.

6.4 Results and Discussion

The simulated linear infrared absorption and 2D IR spectra for the \( \theta \)-states are given in Fig. 6.9. The spectra for the intermediate states, \( S_{\theta}^1 \), and \( S_{\theta}^4 \), are combined to obtain the spectra for state \( I^\theta \). The linear spectra of these
states look similar, except on the left side of the spectra originating from the amide I vibration of the units preceding proline (Y4 and Y9). The peak in this spectral region \((\frac{\omega}{2\pi c} < 1635 \text{ cm}^{-1})\) is more pronounced for \(S_3^\theta\) and \(I^\theta\) than for \(S_2^\theta\).

Figure 6.9: Simulated FTIR and 2D IR spectra for the \(\theta\)-states at 330 K (top) and 300 K (bottom). Contour lines in 2D spectra are plotted between -100% and 100% with a spacing of 10% between two neighboring contours. Red indicates positive and blue indicates negative.
Chapter 6: Temperature-jump relaxation kinetics of a β-hairpin

Figure 6.10: The linear absorption difference spectrum (top) and 2D IR difference spectrum (bottom) between $I^\theta$ and $S_2^\theta$ (left) and between $S_3^\theta$ and $S_2^\theta$ (right) at 330 k.

The general features of the 2D spectra (Fig. 6.9) are similar to those observed before for other β-hairpins [40, 46, 49]. For example, the ridge like cross peak region ($(\omega_1, \omega_3) \sim 1660-1675$ cm$^{-1}$) arises due to excitonic couplings between the amide I oscillators in the peptide. Like in the linear absorption spectra, the significant difference between the 2D IR spectra for the $\theta$-states is reflected in the spectral region for $\omega_1^{2\pi c}$, $\omega_3^{2\pi c} < 1635$ cm$^{-1}$. That is, the amide I vibrational band for the units preceding proline are more red-shifted for $S_3^\theta$ and $I^\theta$ than that for $S_2^\theta$. The difference spectra between these states are presented in Fig. 6.10.

To find out the origin for the spectral differences between the $\theta$-states, we investigated the H-bonding environment between water and Y4 (Y9) analyzing the 330 K MD trajectory. We found that the probability of forming a H-bond between Y4 (Y9) and water is high ($\sim$75%) in $S_3^\theta$ and is very low ($\sim$10%) in state $S_2^\theta$ (Fig: 6.11). These are consequences of the fact that in state $S_2^\theta$ the side chains of both lysines are oriented towards the turns resulting in close distances between these bulky side chains and Y4 (Y9). Therefore, water molecules get little access to these amide I units. In case of state $S_3^\theta$, the side chains of both lysines point away from the turns, resulting in an increases probability that both the Y4 and Y9 units get
Figure 6.11: Histogram for number of H-bonds between carbonyl oxygen of Y4 (Y9) and water in the θ-states.

water exposed. For the intermediate states ($I^θ$) there is a high probability (65%) to find at least one unit water exposed, i.e. Y4 is water exposed in $S^θ_1$ and Y9 is water exposed in $S^θ_4$. The water exposure is responsible for the more red-shifting of the amide I band corresponding to units preceding proline in the spectra for $S^θ_3$ and $I^θ$ relative to that for the $S^θ_2$-spectrum. In the 2D IR studies of an elastin like peptide a similar observation has been reported [222]; the shift of the peak corresponding to the amide unit preceding proline depends on the change in the number of hydrogen bonds between this unit and the peptide (or solvent).

The populations of the θ-states change when a T-jump is performed from 300 K to 330 K and the spectroscopic signatures of these states significantly differ from each other in terms of the amide I vibrations of the units preceding proline. Thus, we choose to track changes in the spectral region corresponding to these amide units, following the T-jump. To determine the spectral changes while the system relaxes to a new equilibrium after the T-jump, we calculated the weighted linear infrared absorption and 2D IR spectra at 330 K as a function of time (Eq. 6.6) considering all θ-states. The weighted spectrum at 300 K is the spectrum before the T-jump and is subtracted from the weighted spectra at 330 K at different time points to obtain the difference spectra (Eq. 6.7). In the linear infrared absorption difference spectra (Fig. 6.12) we observe a gradual growth of the intensity with time in the region between 1610 cm$^{-1}$ and 1630 cm$^{-1}$. In the 2D IR difference spectra (Fig. 6.13) the negative peak (centered at $\omega_{2πc} = 1619$ cm$^{-1}$, $\omega_{2πc} = 1619$ cm$^{-1}$) gains positive intensity and the intensity of the
positive peak (centered at $\frac{\omega_1}{2\pi c} = 1619 \text{ cm}^{-1}$, $\frac{\omega_3}{2\pi c} = 1601 \text{ cm}^{-1}$) decays with time. To display the spectral changes more clearly, slices are extracted from the 2D IR difference spectra (Fig. 6.13) at $\frac{\omega_1}{2\pi c} = 1619 \text{ cm}^{-1}$. The intensity change with time follows a similar trend as was observed for the linear infrared absorption difference spectra (Fig. 6.12). This growth of intensity is attributed mainly to the population increase of state $S^\theta_3$ for which the spectral region corresponding to the amide I units preceding proline is well resolved and the population decrease of state $S^\theta_2$ for which it is not resolved.

To determine quantitatively the time evolution of the intensity growth, the intensity at $\frac{\omega_1}{2\pi c} = 1615 \text{ cm}^{-1}$ and $\frac{\omega_3}{2\pi c} = 1624 \text{ cm}^{-1}$ of the linear infrared absorption difference spectra (Fig. 6.12) for all time points were calculated. For the slices through the 2D IR difference spectra (Fig. 6.13), the intensity at $\frac{\omega_1}{2\pi c} = 1601 \text{ cm}^{-1}$ and $\frac{\omega_3}{2\pi c} = 1619 \text{ cm}^{-1}$ for all time points are also collected. These are depicted in Fig. 6.14 and it is possible to fit a single exponential function to all these curves resulting in the same time scale ($\sim 45 \text{ ps}$) for the intensity evolution. This means that there is one dominant relaxation process, although a combination of multiple exponential process are expected to be observed. If the population evolution of the $\theta$-states (Fig. 6.8b) are fit with an exponential function, the obtained relaxation...
6.4: Results and Discussion

Figure 6.13: **Top Panel**: 2D IR difference spectra following the T-jump; contours plotted between -0.8 and 0.6 with a spacing of 0.06 between neighboring contours. Green dashed lines are drawn to indicate $\frac{\omega_1}{2\pi c} = 1619 \text{ cm}^{-1}$.

**Bottom Panel**: Slices through the 2D IR difference spectra at $\frac{\omega_1}{2\pi c} = 1619 \text{ cm}^{-1}$ with zoomed in regions around 1601 cm$^{-1}$ and 1619 cm$^{-1}$, respectively. The slices are extracted at the same time points as for the linear absorption difference spectra (Fig. 6.12).

Times are 45 ps, 50 ps, and 30 ps, respectively for $S_2^\theta$, $S_3^\theta$, and $I^\theta$. Thus, the relaxation process observed in the spectral changes following the T-jump is dominated by the population decay of $S_2^\theta$. We note that these time scales correspond to the eigenvalues of the transition matrix for the $\theta$-states and...
these are smaller than the rates shown in Fig. 6.8a.

We assumed that no structural changes occur during the heating of the peptide-solvent system in the T-jump experiment and following the T-jump the peptide structure changes. Now the question is, "How significant is the effect of heating on the spectra obtained after the T-jump, compared to the effects from the structural changes after the T-jump?" The effects of heating are as follows. In the 2D IR spectra presented in Fig. 6.9 we see that the diagonal and the overtone peaks are positive and negative, respectively, whereas in the 2D IR difference spectrum calculated at 0 ps after the T-jump (Fig. 6.13) these peaks have interchanged their signs in the region corresponding to the amide units preceding proline. After that, the structural changes are reflected in the spectral changes but the signs remain the same. Thus, the effect of heating is more significant than the effect of the structural changes.

The significant spectral changes due to structural changes occur at hundreds of picoseconds in our simulation (Fig. 6.14), which differs from the existing experimental observation [72] where the significant spectral changes were observed around 140 ns. This is because we have restricted
ourselves to studying the sub-nanosecond side chain dynamics of the cyclic peptide subject a T-jump; therefore, the kinetics of the \( \theta \)-states does not reflect the experimentally [72] determined folding/unfolding time scales (\( \sim 140 \) ns). Considering the H-bonding configurations it is possible to obtain the kinetics to compare with experiment, but with the current force fields we do not find the fully unfolded configuration (mid-strand H-bonds remain intact even at a high temperature), although extensive REMD simulations were performed. In a future study different types of force fields should be considered to access the fully unfolded configuration, as it is well known that the configuration space depends on the choice of force fields [60]. Then, using the T-jump protocol described in this study one can find out what particular conformation changes and kinetics are hidden in the experimentally measured folding/unfolding rates.

To resolve the kinetics of the \( \theta \)-states experimentally, a faster triggering mechanism might be needed. A photo chemical triggering mechanism [224] allowing the observation of ps time scale dynamics may thus be useful.

\section*{6.5 Conclusions}

To summarize, we have developed a protocol to model and simulate a T-jump experiment combined with infrared spectroscopy using the following steps. We have investigated the backbone and side chain structures of GS10 with REMD simulations. The kinetics of the states corresponding to the orientations of the lysine side chains obtained from the MD simulations at 330 K are used to model and simulate the T-jump from an initial temperature 300 K to the final temperature 330 K. The system in equilibrium at 300 K relaxes to a new equilibrium at 330 K following the T-jump. During the relaxation process structural changes occur, i.e. the populations of the states change with time until the system reaches its new equilibrium. This phenomenon is reflected in the changes of the linear infrared absorption and 2D IR difference spectra with time. In this way we have simulated a T-jump experiment combined with the infrared techniques allowing to unravel the connection between the conformational dynamics and the spectral changes following a T-jump and determine the underlying kinetics.

The folding or unfolding rates for this peptide have not been determined in this study because our REMD simulations did not access the fully unfolded configuration. This leads us to conclude that the classical force field used here should be improved in order to obtain the correct configurations.
Once the correct configurations are generated, using our T-jump protocol combined with linear infrared absorption and 2D IR experiments, it will be possible to follow the folding/unfolding dynamics and reveal the related kinetics of proteins and peptides. For example, proteins containing proline could be useful systems to study unfolding dynamics, because proline is generally found in the turn and investigating spectral changes of the spectral region corresponding to the amide I unit preceding proline allows one to focus on the turn conformational changes caused by the T-jump.

6.6 Appendix

We present here the histograms of time intervals for the transitions between the $\theta$-states. The exponential fits to the histograms shown in Fig. 6.15 give the times scales for the transitions. The histograms corresponding to the transitions between $S_1^\theta$ and $S_4^\theta$, between $S_4^\theta$ and $S_1^\theta$, $S_2^\theta$ and $S_3^\theta$, and between $S_3^\theta$ and $S_2^\theta$ are not good enough to make exponential fits, because the number of the corresponding transitions were not sufficient to provide nice histograms (Fig. 6.16). In such case, the average of the time intervals for a particular transition is assigned as the time scale for that particular transition.

Figure 6.15: Histograms of time intervals corresponding to the transitions between the $\theta$-states, for which the exponential fits can be done to obtain the transition times.
Figure 6.16: Histograms of time intervals corresponding to the transitions between the θ-states, for which exponential fits cannot be done.
Chapter 6: Temperature-jump relaxation kinetics of a β-hairpin