Protein-protein interactions in model systems – design, control of catalytic activity and biosensor applications

Johan Rydberg

IFM Chemistry, Division of Organic Chemistry
Linköpings universitet, SE-581 83 Linköping, Sweden

Linköping 2006
Abstract

This thesis describes the design of polypeptides, unordered in the monomeric state but capable of folding into helix-loop-helix motifs and dimerise to form four-helix bundles. The goal of the design was to encode them with the capacity to form dimers highly selectively and the ability to carry out molecular functions in the folded state but not in the unordered state, and thus to establish a molecular link between recognition and function. The 42-residue sequences JR2E and JR2K were both shown by CD spectroscopy to adopt unordered conformations under single solute conditions at pH 7 but to form helical conformations in a 1:1 mixture. Analytical ultracentrifugation showed that JR2E and JR2K formed a clean heterodimer and the dissociation constant $K_d$, measured by CD spectroscopy, was found to be $5 \pm 1 \mu$M. Discrimination was enabled by the incorporation of charged residues at the dimer interface in the helical segments of the helix-loop-helix motif. Glutamic acids were incorporated in JR2E and lysines in JR2K, and charge repulsion prevented the monomeric subunits from forming homodimers. In mixtures, however, highly helical heterodimers were formed. The cooperative transition from unordered conformation to heterodimeric four-helix bundle was exploited in the design of a signal response system by incorporating a reactive site, capable of catalysing the hydrolysis of a m-nitrophenyl ester, into the negatively charged polypeptide. In the unfolded state the functionalised polypeptide was virtually inactive but in the folded state, induced by the interaction with JR2K, the substrate was hydrolysed approximately an order of magnitude more efficiently.

Interactions between the designed polypeptides and a functionalised polythiophene polymer were studied and it was found that the conformation of the polymer was controlled by the polypeptides, largely by electrostatic interactions. The negatively charged JR2E forced the polymer to adopt a planar conformation while the positively charged JR2K induced a more twisted conformation of the polymer. The spectral changes coupled to the conformational transitions of the polymer were used to measure the binding of human Carbonic anhydrase II by JR2E functionalised with a benzenesulphonamide ligand, in demonstration of its use as a tool for high-throughput screening.

JR2E immobilised on gold nanoparticles was shown to form homodimers reversibly under pH control, with affinities large enough to determine the state of aggregation of the gold nanoparticles.
Publications

Papers included in this thesis:

I. Rydberg, J.; Vijayalekshmi, S. and Baltzer, L. Intrinsically unstructured proteins by design - electrostatic interactions can control binding, folding and function of a helix-loop-helix heterodimer. Submitted


Contribution report

Paper I. I carried out essentially all of the experimental work. I provided a preliminary version of the manuscript.

Paper II. I carried out the design and synthesis of the polypeptides and took part in the NMR experiments. I did a minor part of the writing.

Paper III. I did 75% of the experimental work. I provided a preliminary version of the manuscript.

Paper IV. I participated in the design of the polypeptides.
Related paper not included in this thesis:

## Contents

1. Introduction ..............................................................................................................3

2. De novo protein design ..........................................................................................5
   2.1. The design of an α-helix ..............................................................................5
   2.2. The four-helix bundle ................................................................................6
   2.2.1. The design of a helix-loop-helix dimer .................................................7
   2.3. The design of a heterodimer ........................................................................8
   2.3.1. The design of a peptide library ..............................................................9
   2.3.2. Structural characterisation .................................................................11
   2.4. Characterisation of the interaction between JR2E and JR2K .................15
   2.4.1. The salt dependence .............................................................................15
   2.4.2. The pH dependence ............................................................................16
   2.4.3. The state of aggregation ......................................................................17
   2.4.4. The affinity between JR2E and JR2K ..................................................19
   2.4.5. The dynamic nature of the heterodimer ..............................................20

3. The design of larger polypeptide scaffolds ..........................................................21
   3.1. The design and synthesis of JR94E and JR94K .......................................22
   3.2. The pH profile ..........................................................................................23
   3.3. The salt dependence ..................................................................................24
   3.4. The state of aggregation ..........................................................................25
   3.5. The affinity between JR94E and JR94K ....................................................26

4. Using heterodimers as biocatalysts .......................................................................27
   4.1. Designed protein catalysts .......................................................................27
   4.2. Induced catalytic activity by selective protein interactions .......................28
   4.2.1. The functionalisation of JR2E ...............................................................28
   4.2.2. ECC as a catalyst .................................................................................30

5. The interplay between designed polypeptides and a conjugated polymer ............31
   5.1. Properties of a zwitterionic polythiophen derivative ..................................31
   5.2. Conformational transitions of POWT induced by synthetic polypeptides ..32

6. JR2E and POWT as a platform for biosensor applications ..................................37
   6.1. The transformation of JR2E into a binder of HCAII ....................................37
   6.1.1. The synthesis of JRE2-lig .....................................................................37
   6.2. Optimisation of the biosensor .....................................................................38
   6.3. The interactions between the peptide/polymer complex and HCAII ...........39
   6.4. Screening for high-affinity protein ligands .................................................41

7. The use of protein-protein interactions in tailoring properties of hybrid materials .................................................................................................................43

8. Methods ....................................................................................................................45
   8.1. Solid-phase peptide synthesis ....................................................................45
   8.2. High performance liquid chromatography (HPLC) .................................45
   8.3. Matrix assisted laser desorption ionisation time of flight (MALDI-TOF) ..46
   8.4. Circular dichroism (CD) spectroscopy .....................................................47
   8.5. Nuclear magnetic resonance (NMR) spectroscopy ....................................48
   8.6. Fluorescence spectroscopy ........................................................................49
   8.7. Analytical ultracentrifugation ...................................................................49

9. References ..............................................................................................................51

10. Acknowledgments ................................................................................................55
1. Introduction

Protein-protein interactions play key roles in many biological processes, for instance in the formation of hormone-receptor, protease-inhibitor and antibody-antigen complexes that are crucial for the function of the life processes. However, not all protein-protein interactions are beneficial. Some serious diseases e.g. sickle cell anemia\cite{1} and Alzheimer’s disease are caused by undesirable protein-protein interactions\cite{2}. Also the loss of protein-protein interactions can have serious consequences. For example the oxidation of α-1-proteinase, which may be caused by cigarette smoke, prevents the inhibition of elastase that in turn may lead to emphysema\cite{3}.

The above examples illustrate the importance of protein-protein interactions in nature and the serious consequences that might arise when they fail. Much effort has been devoted to increasing our understanding of these interactions at the fundamental level\cite{4}. In addition, the finding that some serious diseases are caused by uncontrolled protein-protein interactions and that key enzymes in pathogens such as HIV depend for their function on being in the right oligomeric state have led to the development of new drug candidates that target protein interfaces\cite{4,5}. However, learning more about protein-protein interactions is not only important in the quest for new drugs. An improved fundamental understanding of protein-protein interactions will perhaps make it possible to design more complex protein structures with multiple functions.

The main purpose of this thesis was to explore how charged residues at the interface between two helix-loop-helix motifs could be used to control the oligomeric state of the polypeptides and to demonstrate some of the many applications that are enabled by heterodimer discrimination. Chapter two and three of this thesis describe the design of heterodimeric polypeptides that are highly dependent on interactions between the subunits to adopt their folded conformations, and the level of selectivity that can be obtained in heterodimer formation. In chapter four one of the subunits in a heterodimer is functionalised with a catalytic site in an attempt to use protein-protein interactions in a signal response system where the functionalised polypeptide is
catalytically active only in the presence of a charge and shape complementary activator. In chapter five and six the protein-protein interaction concept is taken one step further to also include protein-organic polymer interactions and it is shown how these complexes may be used in biosensor applications. In chapter seven a designed polypeptide is immobilised on gold nanoparticles to demonstrate how protein interactions can be used to tailor the optical properties of the nanoparticles.

The results presented in this thesis do not only demonstrate how protein-protein interactions between designed polypeptides might be used in various applications, they also provide some insight into the strength of using solid phase peptide synthesis and de novo designed polypeptide scaffolds in the quest for the fundamental understanding of proteins.
2. De novo protein design

Most native protein adopts only a single three-dimensional structure even though the number of theoretical conformational options of a moderately sized protein is astronomical. To predict which conformation the protein will adopt from an inspection of its primary sequence is not trivial and to determine the relationship between the primary sequence and the tertiary structure de novo protein design has emerged as an attractive approach. In de novo protein design a certain fold is predicted from a primary sequence not copied from any native protein. Structural analysis then shows whether or not that prediction was correct and how far we have come in understanding protein structure. Although the relationship between the primary sequence and the tertiary structure is far from understood, considerable progress has been made in the design of small to moderately sized proteins consisting of a few secondary structure elements like the $\alpha$-helix and the $\beta$-sheet$^{[6-8]}$. When we truly understand how to design proteins from scratch it should be possible to design enzymes, receptors or ion channels perhaps with functionalities and properties not resembling those of any native proteins and useful for biomedical or industrial applications.

2.1. The design of an $\alpha$-helix

The best understood secondary structure element is the $\alpha$-helix, in which the polypeptide backbone forms a right-handed spiral with approximately 3.6 amino acid residues per turn and a distance of about 5.6 Å between each turn of the helix. In order for an $\alpha$-helix to fold, the loss in conformational entropy must be overcome. Several design principles have emerged that may be used to stabilise the folded helix relative to the unfolded peptide. The use of residues with large helix-forming propensities such as alanine, leucine and glutamic acid$^{[9]}$ and also the use of intramolecular salt bridges$^{[10]}$ have proven to be good strategies. In the folded helix hydrogen bonds are formed between the backbone carbonyl oxygens and amide protons of residues four positions apart in the polypeptide chain. At the N-terminal and C-terminal ends of the $\alpha$-helix amide protons and carbonyl oxygens, respectively,
will be left with unsatisfied hydrogen bonds. To overcome this, the side chains of the first and last amino acids are selected from those that have hydrogen bond acceptors and can bind to the backbone amide protons or have hydrogen bond donors that can interact with the backbone carbonyl oxygens\[^{11,12}\]. These residues, e.g. Asn, Asp, Ser, Thr or Gly, are known as N- and C-terminal capping residues. In the α-helix all carbonyl groups of the backbone points in the same direction, leading to a macroscopic dipole moment with a partial positive charge at the N-terminus and a partial negative charge at the C-terminus. Introduction of residues with opposite charges near the helical ends will neutralise the dipole and hence stabilise the α-helix\[^{13}\].

![Diagram](image)

**Figure 1.** *Schematic representation of an α-helix showing some of the interactions that stabilise the folded conformation.*

### 2.2. The four-helix bundle

The naturally occurring four-helix bundle structure\[^{14,15}\] is a motif that has been the target of several *de novo* designs\[^{16-18}\]. The fold consists of four aggregated helices and may be formed from four separate amino acid sequences in a tetrameric structure. The helix segments can also be connected by short loops in helix-loop-helix dimers or in single chain four-helix bundle motifs. In the early days of protein design many of the reported designed four-helix-bundles showed the characteristics of molten globules\[^{16,17,19}\]. In a molten globule there are several different structures in rapid equilibrium resulting in poorly defined melting points and poorly dispersed NMR spectra. The molten globule is often characterised by a high content of secondary
structure. For a four-helix bundle, or any other polypeptide chain, to adopt a unique fold there must be a large free energy gap between the correct fold and every other folded state\textsuperscript{[20]}. Achieving this large free energy gap is not an easy task but successful designs have been reported\textsuperscript{[21-23]}. A fruitful approach in the design of native-like structures has been the incorporation of negative design elements, i.e. using residues that do not contribute to the stability of the correct fold but destabilise alternative folds\textsuperscript{[24]}.

A commonly occurring feature of four-helix bundle proteins is that the helices are amphiphilic with one polar and one hydrophobic side. In the case of a water-soluble protein the hydrophobic faces are not solvent exposed but form a hydrophobic core. The four-helix bundle is conveniently described in terms of the heptad repeat pattern \((a\ b\ c\ d\ e\ f\ g)_n\) (Figure 2). The \(a\) and \(d\) positions form the hydrophobic core while the other positions are exposed to the exterior. The \(b\) and \(e\) positions are located at the dimer interface and control the mode of dimerisation.

\textbf{Figure 2. Helical wheel representation of an antiparallel helix-loop-helix dimer. Helix I and II and helix I’ and II’ are connected by short loops not shown in the figure. The direction of helix I is in to the plane and that of helix II is out of the plane, while the opposite is true for helix I’ and helix II’.

2.2.1. The design of a helix-loop-helix dimer

SA-42 is a designed polypeptide, reported previously, that forms an antiparallel helix-loop-helix homodimer\textsuperscript{[17]}. SA-42 consists of 42 amino acid residues that were selected based on their helix and loop forming propensities. To further stabilise the folded structure capping residues were introduced and charged residues were placed near the N- and C-terminus of each helix to neutralise the helical dipole. Hydrophobic residues were incorporated in the \(a\) and \(d\) positions and charged residues capable of forming
salt bridges were introduced in solvent exposed positions. Folding and dimerisation of a helix-loop-helix motif can lead to a parallel, antiparallel or a bisecting U motif (Figure 3), but structural analysis showed that SA-42 adopted the antiparallel conformation\textsuperscript{17}.

![Figure 3](image)

**Figure 3.** Three different folding options for a helix-loop-helix dimer. From left to right the parallel, the antiparallel and the bisecting U dimers.

### 2.3. The design of a heterodimer

The main objective in the design of a heterodimer model system was that the peptides in the monomeric state should not adopt ordered conformations but have a high propensity for the formation of ordered heterodimers (Figure 4). The four-helix bundle SA-42 was chosen as the template with hydrophobic residues in the $a$ and $d$ positions, the $c$ and $g$ positions available for the engineering of a catalytic site and the $b$ and $e$ positions to be used to control dimerisation at the dimer interface (Figure 2). In order to ensure inhibition of homodimerisation, residues with the same charge were introduced in positions at the dimer interface to cause charge-charge repulsion between the peptides. Charge-charge repulsion within each helix was also expected to contribute to inhibition since distances between charged residues would be shorter in the folded than in the unfolded state. If the repulsive forces between the charged residues are greater than the attractive van der Waals interactions between the hydrophobic residues in the core, then the peptides will not be able to adopt an ordered structure. In a 1:1 mixture of two peptides with opposite charges, however, there will be electrostatic attraction in addition to the van der Waals interactions between the peptides leading to the formation of a heterodimer. It was expected that the discrimination in heterodimer formation could be controlled by electrostatic interactions and that it would depend on the number and positions of charged residues at the dimer interface.
Figure 4. Illustration of the design objective showing a negatively charged polypeptide with a random conformation (left) and a positively charged polypeptide with a random conformation (middle) that does not form homodimers. The electrostatic interactions favour the formation of a helix-loop-helix heterodimer (right).

2.3.1. The design of a peptide library

In order to determine which positions of the polypeptide sequence that were best suited for introduction of charged residues in the search for maximum repulsion, and also how many charged residues that were required for inhibition of dimerisation a small library of 12 peptides was created. The peptides were JR1E to JR6E (the E-peptides) and JR1K to JR6K (the K-peptides) (Table 1). Broadly, the lower number sequences (JR1E, JR1K etc) have charged residues in \( b \) and \( e \) positions, whereas higher number sequences have charged residues in \( b \) and \( e \) positions as well as in \( f \) positions. Odd numbered sequences (JR1E, JR3E etc) have three charged residues at the dimer interface of each helix whereas even numbered sequences have four charged residues in these positions. The highest numbered sequences are combinations of helices from the lower number sequences. A problem in the design of the negatively charged peptides was the poor solubility. To overcome this problem, lysines were placed in \( f \) positions.

The E-peptides have glutamic acids at the dimer interface while the K-peptides have lysines. In JR1E the glutamic acids were placed in positions 6, 10, 13, 28 and 32 which are \( b \) and \( e \) positions and in position 36 that is an \( f \) position. Unpublished results have shown that the phenylalanine in the \( e \) position 35 is important for the stability of the structure and was therefore not modified.

In JR2E the two additional glutamic acids were introduced in positions 17 and 39 (\( b \) positions). The glutamic acids in JR3E were placed in positions 7, 10, 14, 29, 32 and 36 (\( b \) and \( f \) positions). In JR4E the two additional glutamic acids were placed in
positions 17 and 39 (b positions). In JR1E and JR2E there are lysines in the f positions 7 and 29. In JR3E the corresponding lysines were introduced into the c positions 11 and 33. JR5E is a combination of helix I in JR1E with helix II in JR3E and JR6E is a combination of helix I in JR2E with helix II in JR4E. JR1K-JR6K corresponds to JR1E-JR6E but have lysines instead of glutamic acids in the positions described above.

Table 1. The primary sequences of SA-42 and the peptides JR1E-6E and JR1K-6K. Highlighted positions show residues that were varied in the peptide library.

<table>
<thead>
<tr>
<th>SA-42</th>
<th>JR1E</th>
<th>JR2E</th>
<th>JR3E</th>
<th>JR4E</th>
<th>JR5E</th>
<th>JR6E</th>
<th>JR1K</th>
<th>JR2K</th>
<th>JR3K</th>
<th>JR4K</th>
<th>JR5K</th>
<th>JR6K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>Aib</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>Nie</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>6</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>A</td>
<td>E</td>
<td>E</td>
<td>K</td>
<td>K</td>
<td>A</td>
<td>A</td>
<td>K</td>
</tr>
<tr>
<td>7</td>
<td>K</td>
<td>K</td>
<td>E</td>
<td>E</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>8</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>9</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>10</td>
<td>K</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>A</td>
<td>K</td>
<td>K</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>K</td>
<td>K</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>12</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>13</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>A</td>
<td>E</td>
<td>E</td>
<td>K</td>
<td>K</td>
<td>A</td>
<td>A</td>
<td>K</td>
</tr>
<tr>
<td>14</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>15</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>16</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>17</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>18</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>19</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>20</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>21</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>22</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>23</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>24</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>25</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>26</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
</tr>
<tr>
<td>27</td>
<td>N</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>28</td>
<td>Aib</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>K</td>
<td>K</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>29</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>30</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
</tr>
<tr>
<td>31</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>32</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>34</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>E</td>
<td>A</td>
<td>E</td>
<td>A</td>
<td>K</td>
<td>A</td>
<td>K</td>
<td>A</td>
<td>K</td>
</tr>
<tr>
<td>35</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>36</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

10
2.3.2. Structural characterisation

The mean residue ellipticity at 222 nm, $[\Theta]_{222}$, was used to compare the helical content of the peptides$^{[25,26]}$ at pH 5 and at pH 7 (Table 2).

For both the E- and K-peptides at pH 7 there is a clear relationship between the number of charged residues at the dimer interface and the ability to fold, where the peptides with 8 charged residues at the dimer interface are virtually unordered. For the E-peptides at pH 7 the helical content of the peptides with 6 glutamic acids at the helical interface (JR1E, JR3E and JR5E) was higher than expected for unordered conformations, with mean residue ellipticities in the range from -12800 - -20900 deg cm$^2$ dmol$^{-1}$. Apparently, there was not enough charge repulsion to prevent folding and dimerisation of these sequences. JR1E with glutamic acids in $b$ and $e$ positions had a lower helical content than JR3E with glutamic acids in $b$ and $f$ positions. This was to be expected from an inspection of the helical wheel. The interhelical repulsion in JR1E should be larger since the distance between charged residues is shorter. Incorporation of two additional glutamic acids (JR2E, JR4E and JR6E) led to a considerable decrease in the helical content. This was probably due to a combination of inter- and intrahelical repulsion, although the interhelical repulsion is probably of greater importance. If the decrease in helical content only depended on interhelical repulsion, however, one would not expect much difference between JR3E and JR4E. Surprisingly, JR4E is the E-peptide with the second lowest helical content, even though the negatively charged residues are located in $b$ and $f$ positions. The reason why JR4E has a lower helical content than, for example, JR6E might be that the additional glutamic acid in position 17 causes more intrahelical charge repulsion in JR4E than in JR6E since the glutamic acid in JR6E will be partly neutralised by the lysine in position 14. JR4E also has one more negative charge than JR6E that might destabilise the structure. The glutamic acid in position 17 of JR2E will be partly neutralised as in JR6E, but on the other hand JR2E has more unfavourable $b$ and $e$ interhelical interactions than both JR4E and JR6E and therefore a lower helical content.

For the K-peptides at pH 7 the trend is similar to that of the E-peptides except that JR4K, corresponding to JR4E, has a larger helical content than expected from the number of lysines incorporated in the sequence. Here there are no favourable salt bridges in any of the peptides so the reason for the relatively high helical content of JR4K is probably due to less unfavourable interhelical interactions compared to JR1K and JR6K, even though JR4K is more charged than JR1K.
At pH 5 there is a significant increase in the helical content of most of the E-peptides in comparison with those at pH 7. At pH 5 the glutamic acids, with typical side chain pK\textsubscript{a} values of about 4.1 in model peptides, are partly protonated and less charged. Charge-charge repulsion within and between the peptides is therefore less pronounced. The high helical content of the E-peptides at pH 5 compared to the helical content at pH 7 clearly shows that once the polypeptides do not repel each other they are able to fold.

Table 2. The mean residue ellipticities at 222 nm ([\Theta]_{222}/\text{deg cm}^2 \text{dmol}^{-1}) of the peptides JR1E-6E and JR1K-6K at pH 7 and at pH 5, measured at ambient temperature and a peptide concentration of 0.5 mM.

<table>
<thead>
<tr>
<th></th>
<th>[\Theta]_{222} pH 7</th>
<th>Total charge</th>
<th>[\Theta]_{222} pH 7</th>
<th>Total charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR2E</td>
<td>-4300</td>
<td>-5</td>
<td>JR2K</td>
<td>-3600</td>
</tr>
<tr>
<td>JR4E</td>
<td>-6000</td>
<td>-6</td>
<td>JR6K</td>
<td>-3600</td>
</tr>
<tr>
<td>JR6E</td>
<td>-7300</td>
<td>-5</td>
<td>JR1K</td>
<td>-4200</td>
</tr>
<tr>
<td>JR1E</td>
<td>-12800</td>
<td>-3</td>
<td>JR4K</td>
<td>-7300</td>
</tr>
<tr>
<td>JR5E</td>
<td>-13600</td>
<td>-3</td>
<td>JR5K</td>
<td>-8400</td>
</tr>
<tr>
<td>JR3E</td>
<td>-20900</td>
<td>-4</td>
<td>JR3K</td>
<td>-14400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>[\Theta]_{222} pH 5</th>
<th>Total charge</th>
<th>[\Theta]_{222} pH 5</th>
<th>Total charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR6E</td>
<td>-19500</td>
<td>-1</td>
<td>JR2K</td>
<td>-2600</td>
</tr>
<tr>
<td>JR4E</td>
<td>-20000</td>
<td>-2</td>
<td>JR6K</td>
<td>-3400</td>
</tr>
<tr>
<td>JR5E</td>
<td>-20800</td>
<td>0</td>
<td>JR1K</td>
<td>-3500</td>
</tr>
<tr>
<td>JR2E</td>
<td>-23100</td>
<td>-1</td>
<td>JR5K</td>
<td>-3500</td>
</tr>
<tr>
<td>JR1E</td>
<td>-25100</td>
<td>0</td>
<td>JR4K</td>
<td>-4100</td>
</tr>
<tr>
<td>JR3E</td>
<td>-26700</td>
<td>-1</td>
<td>JR3K</td>
<td>-5000</td>
</tr>
</tbody>
</table>

The K-peptides are more charged at pH 5 than at pH 7, due to protonation of the histidine in position 15. The effect of the additional positive charge on the helical contents of JR1K, JR2K and JR6K is insignificant as expected since these peptides do not adopt any ordered conformation at pH 7. But, for JR3K, JR4K and JR5K the
additional charge reduces at least partly helical structures to largely unordered conformations.

In order to investigate the selectivity in heterodimer formation all combinations of E-peptides with K-peptides were studied in 1:1 mixtures at pH 7 and the mean residue ellipticities determined (Table 3). The helical contents were high in each mixture, in the range from $-21800 \text{ - } -14000 \text{ deg cm}^2 \text{ dmol}^{-1}$. The highest helical contents were observed for combinations of peptides with charged residues in \( b \) and \( e \) positions, with the exception of the combination of JR1E and JR3K where JR3K has the charged residues in \( b \) and \( f \) positions. In this combination the helical content of each peptide was high, $[\Theta]_{222} = -12800$ and $-14400 \text{ deg cm}^2 \text{ dmol}^{-1}$, for JR1E and JR3K, respectively, showing that a high helical content may be due to homodimer formation as well as to heterodimer formation. To clearly establish the extent of heterodimerisation the negative value of the difference between the ellipticity of the combined peptides and the average ellipticity of the single peptides, the $\Delta$ value, was used instead (Table 3). Solutions of combinations of polypeptides with high propensities for homodimer formation, for example that of JR1E and JR3K, will give rise to large negative mean residue ellipticities although perhaps not for the reason that heterodimers are formed. The same result is expected for solutions of combinations of polypeptides with high propensities for heterodimer formation. The $\Delta$ value will discriminate between the two situations. Small $\Delta$ values arise from combinations of peptides that form large populations of homodimers and probably small populations of heterodimers, whereas large $\Delta$ values result from peptides that form large populations of heterodimers but small or nonexistent populations of homodimers. The largest $\Delta$ values were observed for the combinations JR2E/2K and JR2E/1K. None of these polypeptides showed significant populations of ordered conformations under single solute conditions and most of their charged residues were introduced in \( b \) and \( e \) positions. JR2E/2K is therefore a clean heterodimer.

The results show that charged residues located in \( b \) and \( e \) positions efficiently prevent homodimerisation by charge repulsion, and favour heterodimerisation, provided that there are four in each helix. In many combinations there is an increase in the helical content, but not to the same extent as in JR2E/JR2K or in JR2E/JR1K. In the mixtures where one or both of the single peptides are well folded e.g. JR1E and JR3K there is probably a combination of homodimers and heterodimers.
Table 3. The mean residue ellipticities at 222 nm, $[\theta]_{222}$ (deg cm$^2$ dmol$^{-1}$), of all possible 1:1 combinations of E and K peptides at pH 7. The concentration of each peptide was 0.25 mM and the total peptide concentration was 0.5 mM in each sample. The $\Delta$ value is the difference between the measured mean residue ellipticity of the 1:1 mixture and the average of the mean residue ellipticities of the single peptides at pH 7.

<table>
<thead>
<tr>
<th>Combination</th>
<th>$[\theta]_{222}$ pH 7</th>
<th>$\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR1E/3K</td>
<td>-21800</td>
<td></td>
</tr>
<tr>
<td>JR2E/1K</td>
<td>-21100</td>
<td></td>
</tr>
<tr>
<td>JR2E/2K</td>
<td>-21000</td>
<td></td>
</tr>
<tr>
<td>JR1E/2K</td>
<td>-20800</td>
<td></td>
</tr>
<tr>
<td>JR6E/5K</td>
<td>-19800</td>
<td></td>
</tr>
<tr>
<td>JR3E/3K</td>
<td>-19700</td>
<td></td>
</tr>
<tr>
<td>JR1E/1K</td>
<td>-19700</td>
<td></td>
</tr>
<tr>
<td>JR3E/1K</td>
<td>-19300</td>
<td></td>
</tr>
<tr>
<td>JR2E/3K</td>
<td>-19000</td>
<td></td>
</tr>
<tr>
<td>JR6E/6K</td>
<td>-18900</td>
<td></td>
</tr>
<tr>
<td>JR4E/6K</td>
<td>-18900</td>
<td></td>
</tr>
<tr>
<td>JR5E/5K</td>
<td>-18900</td>
<td></td>
</tr>
<tr>
<td>JR6E/2K</td>
<td>-18800</td>
<td></td>
</tr>
<tr>
<td>JR4E/5K</td>
<td>-18600</td>
<td></td>
</tr>
<tr>
<td>JR3E/2K</td>
<td>-18500</td>
<td></td>
</tr>
<tr>
<td>JR6E/4K</td>
<td>-18300</td>
<td></td>
</tr>
<tr>
<td>JR6E/1K</td>
<td>-18200</td>
<td></td>
</tr>
<tr>
<td>JR1E/6K</td>
<td>-18200</td>
<td></td>
</tr>
<tr>
<td>JR5E/6K</td>
<td>-18200</td>
<td></td>
</tr>
<tr>
<td>JR5E/4K</td>
<td>-17900</td>
<td></td>
</tr>
<tr>
<td>JR6E/3K</td>
<td>-17500</td>
<td></td>
</tr>
<tr>
<td>JR5E/2K</td>
<td>-17500</td>
<td></td>
</tr>
<tr>
<td>JR5E/1K</td>
<td>-17500</td>
<td></td>
</tr>
<tr>
<td>JR1E/4K</td>
<td>-17400</td>
<td></td>
</tr>
<tr>
<td>JR5E/3K</td>
<td>-17200</td>
<td></td>
</tr>
<tr>
<td>JR1E/5K</td>
<td>-17000</td>
<td></td>
</tr>
<tr>
<td>JR4E/2K</td>
<td>-16900</td>
<td></td>
</tr>
<tr>
<td>JR4E/1K</td>
<td>-16600</td>
<td></td>
</tr>
<tr>
<td>JR4E/3K</td>
<td>-16500</td>
<td></td>
</tr>
<tr>
<td>JR2E/5K</td>
<td>-16300</td>
<td></td>
</tr>
<tr>
<td>JR3E/5K</td>
<td>-16100</td>
<td></td>
</tr>
<tr>
<td>JR4E/4K</td>
<td>-15900</td>
<td></td>
</tr>
<tr>
<td>JR3E/6K</td>
<td>-15800</td>
<td></td>
</tr>
<tr>
<td>JR3E/4K</td>
<td>-15600</td>
<td></td>
</tr>
<tr>
<td>JR2E/4K</td>
<td>-14800</td>
<td></td>
</tr>
<tr>
<td>JR2E/6K</td>
<td>-14000</td>
<td></td>
</tr>
</tbody>
</table>
2.4. Characterisation of the interaction between JR2E and JR2K

The peptides JR2E and JR2K were selected for further structural characterisation since they were unordered under single solute conditions at pH 7 and demonstrated a high propensity for heterodimer formation.

2.4.1. The salt dependence

To verify the electrostatic nature of the discrimination mechanism the mean residue ellipticities at 222 nm of JR2E and JR2K were measured as a function of NaCl concentration at pH 7 (Figure 5). If JR2E and JR2K do not homodimerise at pH 7 because of charge-charge repulsion it should be possible to shield the charged residues, and induce helix formation by increasing the salt concentration. In contrast, a mixture of JR2E and JR2K should not be affected by salt. This was born out by experiment. At high salt concentrations JR2E and JR2K adopted helical conformations, although they were unordered at low salt, and at a NaCl concentration of 2 M the helical contents of the single peptides were virtually the same as that of the 1:1 mixture. In addition, the helical content of the heterodimer was not affected much by an increased NaCl concentration. These observations provide strong support for the conclusion that discrimination is under electrostatic control and that the charged residues act as negative design elements i.e. they do not contribute much to the stability of the overall fold but they discriminate between alternative folding options. The driving force for structure formation is most likely the hydrophobic interactions in the core.

![Figure 5](image.png)

**Figure 5.** The mean residue ellipticity at 222 nm as a function of [NaCl] for (♦) 50 µM JR2E, (◊) 50 µM JR2K and (●) 25 µM JR2E mixed with 25 µM JR2K in 20 mM Bis-Tris buffer at pH 7.
2.4.2. The pH dependence

To further probe the influence on peptide conformations of the charges of acidic and basic residues, the helical content was measured in the pH interval from 2 to 13 (Figure 6). At a pH below 5 the helical content of JR2E was well developed, clearly due to the lack of charge-charge repulsion within and between the peptides. As the pH was increased the glutamic acids were increasingly unprotonated and negatively charged. Above pH 5 there was a sharp decrease in the helical content as a function of pH and above pH 7 JR2E was largely unordered. In spite of the fact that the pK$_a$ of glutamic acid in a model peptide is 4.1 the helical content of JR2E was unexpectedly high at pH 5. An explanation might be that the pK$_a$ values of the glutamic acids were increased due to the high density of negative charge and that not all of the glutamic acids have to be protonated to allow an ordered conformation to be formed. In the case of JR1E it was found that three glutamic acids in each helix was not enough to generate the repulsive forces required to completely disrupt the structure. JR2E has two more glutamic acids than JR1E and if one or both of these residues are protonated the attractive van der Waals interactions outpower the repulsive forces. This subtle balance between attraction and repulsion may partly be the reason for the observed pH dependence.

![Figure 6](image-url)

**Figure 6.** The mean residue ellipticity at 222 nm as a function of pH for (♦) 50 µM JR2E, (◊) 50 µM JR2K and (●) 25 µM JR2E and 25 µM JR2K.

For JR2K the trend was reversed with a random conformation adopted at low pH where the lysines are protonated. Above pH 9 the lysine residues were increasingly unprotonated as a function of pH and the charge-charge repulsion that inhibits helix formation was decreased. At pH 11 and above the structure was fully developed. The
pH profile of the mixed species indicates that at pH 4 and below where the glutamic acids are protonated to a large degree and the lysines are fully protonated, the mixture contains a single folded species, the homodimer of JR2E. The observation that the mean residue ellipticity of the 1:1 mixture at low pH was approximately half of that of JR2E under single solute conditions is in agreement with this conclusion since the concentration of JR2E in the mixture is half of that under single solute conditions. If there was equal probability for heterodimer and homodimer formation at a pH where there is little charge-charge repulsion within the homodimer of JR2E, then there would be no reason to expect a decrease in the helical content by a factor of two. Above pH 4 where the glutamic acids are increasingly unprotonated, the helical content of the mixture was increased due to heterodimer formation. It is interesting to note that the helical content of the mixture is increased at a lower pH than the pH where the helical content of JR2E is decreased, showing that the heterodimer is somewhat favoured over the homodimer at a pH where glutamic acid side chains of JR2E are partly protonated. Between pH 7 and pH 10 the helical content of the mixture reaches a plateau where the predominant species is the heterodimer because no homodimers are formed. Above pH 10 where the lysines are increasingly unprotonated the helical content of the mixture is less than that of JR2K. Also at the high end of the pH scale there are indications that the heterodimer is more favoured than the homodimer. At pH 11 there is only a slight difference between the helical content at this pH and the concentration of JR2K in the mixture is half of that under single solute conditions, a major contribution to the helical content in the mixture must come from the heterodimer. Only at pH 12 are there indications of a single homodimer.

2.4.3. The state of aggregation

The CD spectra of JR2E, JR2K and the 1:1 mixture of the two clearly indicate that the folded species is a heteromer (Figure 7) but they do no not show if it is a dimer, trimer or other oligomeric species. SA-42, the template peptide, was shown to fold into a helix-loop-helix dimer but JR2E and JR2K differ significantly from the template peptide and it is not safe to assume that they too will form a dimer. To determine the state of aggregation of JR2E and JR2K they were analysed by analytical ultracentrifugation. The results show that a dimer is formed in a 1:1 mixture of JR2E and JR2K at 0.1 mM concentration. The sum of the calculated molecular weights for the peptides was 9154 g/mol and the experimentally determined value was 9177 g/mol (Table 4). No evidence for further aggregation was obtained. The
experimentally determined molecular weights of JR2E and JR2K differed from the theoretical values.

![CD spectra](image)

**Figure 7.** CD spectra of 0.5 mM JR2E (○), 0.5 mM JR2K (▲) and of 0.5 mM JR2E/JR2K (♦) in 20 mM Bis-Tris at pH 7.

The experimentally determined molecular weight of JR2E was 4256 g/mol whereas the theoretical weight was 4580 g/mol, and the experimentally determined molecular weight of JR2K was 3950 g/mol as compared to the theoretical value of 4573 g/mol. For both JR2E and JR2K the experimentally determined molecular weights were less than the theoretical. This is expected due to electrostatic repulsion. The sediment action of JR2E is opposed by the repulsive forces between the charged molecules and as a result the apparent molecule weight is less than the theoretical[27]. In the mixture of JR2E and JR2K the charges are neutralised, the system becomes more ideal and the experimental molecular weight agrees better with the theoretical value.

**Table 4.** Molecular weights of JR2E, JR2K and JR2E/JR2K determined by analytical ultracentrifugation. The peptide concentrations in all samples were 0.1 mM in 20 mM Bis-Tris pH 7 with 200 mM NaCl added.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Experimental MW</th>
<th>Theoretic MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR2E</td>
<td>4256 g/mol</td>
<td>4580 g/mol</td>
</tr>
<tr>
<td>JR2K</td>
<td>3950 g/mol</td>
<td>4573 g/mol</td>
</tr>
<tr>
<td>JR2E/JR2K</td>
<td>9177 g/mol</td>
<td>9153 g/mol</td>
</tr>
</tbody>
</table>
2.4.4. The affinity between JR2E and JR2K

In a bimolecular complex formed from molecules A and B,

\[ A + B \rightleftharpoons AB, \]

the dissociation constant \( K_d \) is

\[ K_d = \frac{[A][B]}{[AB]}, \]

where [A] and [B] are the concentrations of free A and B and [AB] is the concentration of the bimolecular complex at equilibrium. A high affinity between A and B is characterised by a low value of \( K_d \) and protein-protein interactions where \( K_d \leq 10^{-8} \text{ M} \) are roughly considered to be high-affinity interactions\(^{[28]}\).

![Figure 8](image)

**Figure 8.** Left, the concentration of folded JR2E as a function of total concentration of JR2K. Right, the concentration of folded JR2K as a function of total concentration of JR2E. The solid lines represent the fit of a function describing the dissociation of a bimolecular complex. A \( K_d \) value of 4 \( \mu \text{M} \) was obtained when JR2E was titrated with JR2K and 6 \( \mu \text{M} \) was obtained when JR2K was titrated with JR2E.

The affinity between JR2E and JR2K was determined from the increase in helical content measured by CD when either a sample of JR2E was titrated with JR2K or vice versa. The percentage of the helical content at each peptide concentration was estimated by using the algorithm K2D\(^{[29]}\). This fraction was then multiplied with the total peptide concentration and divided by two to give the concentration of one of the folded peptide species. An equation describing the dissociation of a bimolecular complex was fitted to the experimental data (Figure 8). The dissociation constant
when JR2E was titrated with JR2K was found to be 4 µM and when JR2K was titrated with JR2E the dissociation constant was found to be 6 µM. A $K_d$ value in the µM range indicates that the forces between the two polypeptides are strong although not a high-affinity interaction.

2.4.5. The dynamic nature of the heterodimer

The CD spectrum showed that the secondary structure of the heterodimer formed from JR2E and JR2K at pH 7 was well developed, but CD spectroscopy can not provide information about whether the structure is native-like or has the characteristics of a molten globule. This can be determined by NMR spectroscopy since the chemical shifts and the line widths of $^1$H NMR spectra are very sensitive to the molecular environment of each proton$^{[30]}$. In the spectra of JR2E and JR2K (Figure 9) the resonances are narrow and the shift dispersion is poor which is typical of poorly ordered polypeptides and due to the formation of several different conformations that exchange at a rate that is fast on the NMR time scale$^{[31]}$. In the spectrum of the mixture of JR2E and JR2K the lines are broader and the chemical shifts are more dispersed indicating a more ordered conformation but not to the level of a native-like protein. The spectrum is typical of a molten globule like ensemble of structures with similar free energies and folds.

![Figure 9. 600 MHz $^1$H NMR spectra of JR2E (upper left), JR2K (upper right) and JR2E/JR2K (lower left) at pH 7 and 298K.](image)
3. The design of larger polypeptide scaffolds

The four-helix-bundle motif has proven to be very useful as a scaffold for several applications such as biosensing\cite{32} and catalysis\cite{33, 34}. However, most naturally occurring proteins form much larger and more complex structures that for example can accommodate cavities for binding and catalysis. The construction of proteins from scratch that can fold into structures large enough to form cavities and active sites for catalysis is currently the biggest challenge in de novo protein design\cite{35}.

In an approach to design a scaffold that could accommodate a cavity or cleft the same concept that had been used in the design of JR2E and JR2K was applied. Two negatively charged helix-loop-helix motifs were connected by a loop, and two positively charged helix-loop-helix motifs were connected by a loop. Lessons from the design of the smaller systems suggested that these sequences would not fold on themselves but favour the formation of a heterodimer (Figure 10).

![Figure 10](image.png)

**Figure 10.** Schematic representation of a possible eight-helical bundle structure formed from one negatively charged sequence capable of folding into four helical segments and one positively charged sequence capable of folding into four helical segments. Due to electrostatic repulsion the sequences were expected to remain unordered under single solute conditions but fold into an eight-helix bundle motif in a 1:1 mixture.
3.1. The design and synthesis of JR94E and JR94K

The sequence of JR94E is essentially two sequences of JR2E connected by a loop composed largely of glycine residues (Figure 11). In an attempt to create a cavity in the central core, arginines were incorporated at positions 11, 15, 30 and 33 in the second JR2E sequence. To avoid excessive charge repulsion at the interface between the four-helix bundle subunits the lysines in positions 7, 14 and 29 of the second JR2E sequence were replaced by alanines. The sequence of JR94K is related to that of JR2K in the same way that JR94E is related to that of JR2E with the difference that the interface residues and accompanying modifications were located in the first JR2K sequence (Figure 11).

![Figure 11. The primary sequences of JR94E (left) and JR94K (right).](image)

Since the yields in solid-phase peptide synthesis decrease with sequence length it is difficult to synthesise polypeptides with more than 60-80 residues. The strategy in the synthesis of JR94E and JR94K was to divide each sequence into two parts and then use native chemical ligation\[36] to couple the polypeptide chains (Scheme 1).
3.2. The pH profile

The pH dependence of the mean residue ellipticity of JR94E (Figure 12) was found to be similar but not identical to that of JR2E (Figure 6). As for JR2E, JR94E was shown to adopt an unordered conformation at high pH due to deprotonation of the glutamic acids. However, JR94E adopts a helical conformation just below pH 8 whereas JR2E does not fold until below pH 7. The pH profile of JR94K also differs from the pH profile of JR2K. JR94K adopts a helical conformation just above pH 8 whereas JR2K does not fold until above pH 9. For both JR94E and JR94K, the charge repulsion mechanism was less effective compared to JR2E and JR2K, perhaps because the larger peptides, in contrast to the smaller ones, can access more helical folding options, where the repulsive forces are not that pronounced. The pH dependence of the 1:1 mixture of JR94E and JR94K is similar to that of the mixture of JR2E and JR2K at low pH. As for the mixture of JR2E and JR2K, the helical content of the mixture of JR94E and JR94K increases just above pH 4. Since JR94K does not contribute to the helical content at a pH below pH 8 and since JR94E does not unfold.
until above pH 6 it is most likely that JR94K interacts with JR94E to form heterodimers. Above pH 8 the pH profiles of JR2E/JR2K and JR94E/JR94K are dissimilar. While the helical content of the mixture of JR2E and JR2K remains constant until pH 11 the helical content of the mixture of JR94E and JR94K decreases above pH 8, the pH where JR94K adopts a helical conformation under single solute conditions. It seems that under conditions where JR94K folds the homodimeric state is favoured over the heterodimeric state. In both systems discrimination in dimer formation was observed, although the windows where the peptides selectively formed homodimers and heterodimers were more narrow for the large polypeptides.

![Graph showing the mean residue ellipticity at 222 nm as a function of pH for different concentrations of JR94E and JR94K.]

**Figure 12.** The mean residue ellipticity at 222 nm as a function of pH for (♦) 50 µM JR94E, (◊) 50 µM JR94K and (•) 25 µM JR94E and 25 µM JR94K.

### 3.3. The salt dependence

As for the JR2E/JR2K system the pH titration of JR94E/JR94K showed that the charged residues play an important role in determining whether the polypeptides will fold. Again, this was verified by the measurement of the dependence of the mean residue ellipticity on the concentration of sodium chloride (Figure 13). JR94E and JR94K under single solute solutions were more sensitive to an increased concentration of salt than the smaller peptides JR2E and JR2K (Figure 5). This is not surprising since according to the pH titration curves for JR94E and JR94K at pH 8 the repulsive forces for both peptides barely prevent them from folding (Figure 12). In the mixture of JR94E and JR94K the conformations are only marginally affected by salt. The experimental evidence does not support any conclusion to the effect that the helical content of the mixture is exclusively due to heterodimer formation at high salt
concentration since the helical contents of the single subunits are virtually equal at each concentration of sodium chloride. There may well be a mixture of homo- and heteromeric species.

Figure 13. The mean residue ellipticity at 222 nm of (◊) 50 µM JR2E, (♦) 50 µM JR2K and (●) 25 µM JR2E and 25 µM JR2K as a function of [NaCl] in 20 mM Trizma pH 8.

3.4. The state of aggregation

A 1:1 mixture of JR94E and JR94K folds into a heterodimer at 0.1 mM concentration according to analytical ultracentrifugation. The molecular weight of the dominant species in a 1:1 mixture of the subunits was found to be 19441 g/mol compared to 19705 g/mol, the sum of the theoretical weights of JR94E and JR94K (Table 5). As in the case of JR2E the experimentally determined molecular weight of JR94E was less than the theoretical value, 3849 g/mol compared to 9860 g/mol. This may again be explained by the negative charges in JR94E leading to an apparent molecular weight that is lower due to the electrostatic repulsion[27]. The reasons why the experimentally determined molecular weight differs more from the theoretical for JR94E than for JR2E may be that the larger polypeptide is more highly charged and that the sample contains less salt. What is more surprising is that JR94K seems to act as a “normal” protein. The molecular weight of JR94K was experimentally determined to be 9629 g/mol compared to the theoretical weight of 9845 g/mol.
Table 5. Molecular weights of JR94E, JR94K and JR94E/JR94K determined by analytical ultracentrifugation. The peptide concentrations in all samples were 0.1 mM in 20 mM Trizma at pH 8 containing 100 mM NaCl.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Experimental MW</th>
<th>Theoretical MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR94E</td>
<td>3849 g/mol</td>
<td>9860 g/mol</td>
</tr>
<tr>
<td>JR94K</td>
<td>9629 g/mol</td>
<td>9845 g/mol</td>
</tr>
<tr>
<td>JR94E/JR94K</td>
<td>19441 g/mol</td>
<td>19705 g/mol</td>
</tr>
</tbody>
</table>

3.5. The affinity between JR94E and JR94K

The dissociation constant of the dimer formed from JR94E and JR94K was estimated to be 1 µM from titration of JR94E with JR94K (Figure 14). This is on the same order of magnitude as that of the complex of JR2E and JR2K. The somewhat higher affinity between JR94E and JR94K compared to JR2E and JR2K may come from interactions between the two four-helix bundles held together by the two loops in the larger complex. The reason why the affinity is not even higher is probably that these surfaces are not optimised for binding and that the arginines will cause some repulsion. The expected entropic advantage in binding two subunits rather than one is obviously offset by the less than perfect fit but most likely also by residual dynamics in the heterodimer.

Figure 14. Titration of a 10 µM solution of JR94E with JR94K. The solid line represents the fit of a function describing the dissociation of a bimolecular complex with a $K_d$ value of 1 µM.
4. Using heterodimers as biocatalysts

In every living organism there is a certain class of molecules, the enzymes, which govern the flow of chemical substances to the benefit of the organism. Almost every enzymes known are proteins, although some RNA molecules have been found to be catalytically active\cite{37}. Some of the most noticeable characteristics of the enzymes are their specificity and catalytic power. Many of the reactions that readily take place in our bodies would not happen without the enzymes. For example in one of the steps of the citric acid cycle, fumarate is converted to (S)-malate by the enzyme fumarase with a turnover number ($k_{cat}$) of 880 s$^{-1}$, the rate constant for the nonenzymatic reaction under similar condition is $2.5 \times 10^{-13}$ s$^{-1}$ which corresponds to a $t_{1/2}$ of 730 000 years\cite{38,39}. Another extreme example is the decarboxylation of orotidylate to uridylate (UMP) in the last step of pyrimidine biosynthesis catalysed by the enzyme OMP decarboxylase with a turnover number ($k_{cat}$) of 39 s$^{-1}$. The rate constant for the nonenzymatic reaction is estimated to be $2.8 \times 10^{-16}$ s$^{-1}$ which correspond to a $t_{1/2}$ of 78 million years\cite{40}. The catalysis takes place in the active site where the substrate, or substrates, is precisely aligned for optimal bond making or breaking and the rate enhancement of the chemical reaction occurs due to the ability of the enzyme to stabilise the transition state.

4.1. Designed protein catalysts

Designing catalytically active proteins is a challenging task since it requires the precise binding and orientation of substrates, transition states and intermediates adjacent to the catalytic machinery. The rational design of polypeptide catalysts is an important approach to increasing our understanding of catalysis and catalyst design. One of the earliest examples of a designed catalytic polypeptide was ‘Oxaldie 1’\cite{41} a 14-residue helix that catalysed the decarboxylation of oxaloacetate through Schiff base formation between the substrate and the N-terminus of the peptide that had a depressed pK$_a$ value. Since then several designed polypeptides have been reported targeting a variety of chemical reactions\cite{42-44}. Another approach is to use a naturally occurring protein and mutate a subset of residues to introduce a catalytic function not
present in nature. This has, for example, been successfully achieved in the reengineering of a glutathione transferase to form a thiolester hydrolase by rational design\(^{[45]}\). Currently the most promising approach to enzyme design is based on computational methods, an impressive example being the conversion of a ribose-binding protein into a triose phosphate isomerase\(^{[46]}\).

### 4.2. Induced catalytic activity by selective protein interactions

Some proteins involved in functional networks are unordered or have large unstructured regions and are inactive until recognised and bound by a specific biomolecular partner. The recognition process results in a disorder-to-order folding transition that activates function\(^{[47, 48]}\). In a similar way we used the highly cooperative folding of JR2E and JR2K to couple recognition and binding to catalytic function.

#### 4.2.1. The functionalisation of JR2E

In order to design a polypeptide catalyst that was dormant in the resting state but activated upon recognition and binding of another biomolecule, the polypeptide JR2E was modified by introducing a catalytic site for ester hydrolysis. The catalytic site was introduced in a way so that the components of the catalytic machinery would be separated in the unfolded state but in close proximity upon formation of the helix-loop-helix motif. In the modified version of JR2E, JR2E\(_{\text{cat}}\) (Figure 15), histidines that may act as nucleophilic, general acid and/or general base catalysts were incorporated in positions 11 and 30. Arginines were introduced in positions 14, 15, 33 and 34 due to their ability to stabilise negatively charged intermediates. This catalytic site, when incorporated in a 42-residue polypeptide based on SA-42, has previously been shown to catalyse the hydrolysis of the RNA analogue 2-hydroxypropyl p-nitrophenyl phosphate (HPNP)\(^{[49]}\) (Figure 16).

![Figure 15. The amino acid sequences of JR2E\(_{\text{cat}}\) (left) and ECC (right).](image)

28
Unfortunately the changes in JR2E\textsubscript{cat} prevented the peptide from folding into a heterodimeric four-helix bundle (Figure 17). It also lost its ability to form a homodimer at low pH where all the glutamic acids are protonated and uncharged. The reason for the inability of JR2E\textsubscript{cat} to fold was probably that the incorporation of the reactive site introduced too many positive charges.

\textbf{Figure 16.} Chemical structures of (left) HPNP and (right) mNPS.

In the second version of the catalyst, ECC (Figure 15), which more resembles the template peptide SA-42 some changes were made in an attempt to increase the helix stability. First the N- and C-terminals were acetylated and amidated, respectively, to prevent unfavourable charge interaction with the helix dipole. Second the branched leucines in most of the \textit{d} positions were replaced by the more helix stabilising unbranched unnatural amino acid norleucine\textsuperscript{[50]} and third, the lysines in positions 7 and 29 were replaced by alanine and glutamic acid, respectively, to prevent unfavourable charge interactions with the catalytic site. This polypeptide showed the desired folding properties in that it was unordered alone but folded into a helical conformation in the presence of JR2K (Figure 17).

\textbf{Figure 17.} The CD spectrum of (left) 0.2 mM JR2E\textsubscript{cat} (\textsuperscript{V}), 0.2 mM JR2K (\textsuperscript{Δ}) and 0.2 mM JR2E\textsubscript{cat}/JR2K (\textsuperscript{◊}) and (right) 0.2 mM ECC (\textsuperscript{▼}), 0.2 mM JR2K (\textsuperscript{▲}) and 0.2 mM ECC/JR2K (\textsuperscript{♦}) in 20 mM Bis-Tris at pH 7.
4.2.2. ECC as a catalyst

The transphosphorylation reaction is difficult to catalyse since the spontaneous hydrolysis of phosphodiester bonds has a half life of approximately 130 000 years\(^{[51]}\).

In an attempt to catalyse the transphosphorylation of HPNP no rate enhancement was observed whether ECC was folded or not. A reason for the lack of catalytic activity of ECC might be that even in the folded state the active site populates too many conformations and that the alignment of the substrate relative to the active site residues required for the hydrolysis is not precise enough. However, upon changing the substrate to a m-substituted nitrophenyl ester (4-sulfamoyl-benzoyle amino-acetic acid 3-nitrophenyl ester, mNPS, Figure 16) ECC was found to be catalytically active. In separate experiments, at peptide concentrations of 50 µM and substrate concentrations of 40 µM at pH 7.0, it was shown that the peptides ECC and JR2K exhibited some catalytic activity although they adopted unordered conformations. A 1:1 mixture of ECC and JR2K, in which a heterodimer was formed, was however considerably more active (Figure 18). ECC in the folded state was eight times more efficient than in the unfolded state. Rate enhancements were calculated from initial rates after subtraction of spontaneous hydrolysis and compensation for differences in concentration. The reaction solution of ECC and JR2K was investigated by MALDI-TOF mass spectrometry and no signs of covalent modification of any of the polypeptides were observed. This shows that molecular recognition in designed polypeptides can trigger function by discriminating between different conformations.

![Figure 18. Initial rates of ester hydrolysis in the presence of (●) 25 µM ECC, 25 µM JR2K and 40 µM mNPS, (▲) 50 µM ECC and 40 µM mNPS, (▼) 50 µM JR2K and 40 µM mNPS and (○) 40 µM mNPS in 20 mM Bis-Tris at pH 7.](image)
5. The interplay between designed polypeptides and a conjugated polymer

Polymers are macromolecules that are built up from several repeating units linked by covalent bonds. Many polymers exist in nature e.g. glycogen, that is built from repeating glucose molecules and serves to store energy in living cells and, of course, proteins built from repeating amino acid units that carry out a variety of biological functions. There are also many synthetic polymers that serve us in our daily life e.g. polyethylene, from ethylene molecules, found in packages and bottles and Teflon, from tetrafluoroethylene molecules, that serves as coatings in many pots and saucepans. Polymers with alternating single and double carbon bonds are called conjugated polymers and have received considerable attention due to their optical properties. Examples of conjugated polymers are polyacetylene and polythiophene.

5.1. Properties of a zwitterionic polythiophen derivative

POWT is a polymer synthesised from chiral 3-substituted thiophene monomers that carry an amino acid function (Figure 19). The side chains of POWT may adopt syn or anti conformations where the anti conformation is somewhat more stable for steric reasons (Figure 19). The optical properties of POWT are due to the variable π-electron overlap along the polymer backbone and are under conformational control. Maximum overlap is achieved in the planar anti conformation and the overlap decreases with increasing deviation from planarity of neighbouring thiophene residues. When the amino acid functionality in POWT is protonated or deprotonated, compared to the zwitterionic condition at pH 5-6, there is a transition from coil-to-rod (nonplanar to planar) conformation resulting in a red shift in the absorption and emission spectra[52]. At low pH the planar conformation is stable but at high pH the polymer chains tend to aggregate and precipitate[52].
Figure 19. Schematic drawings of two different conformations of POWT. Left the syn conformation, right the anti conformation.

5.2. Conformational transitions of POWT induced by synthetic polypeptides

Highly charged polypeptides with several hydrophobic groups were expected to interact strongly with POWT suggesting that self-assembled biosensing systems could be developed and used in biotechnical applications. The addition of JR2E to a solution of POWT in aqueous solution at pH 7.4 strongly affected the fluorescence emission spectrum of POWT (Figure 20). The small difference in emission intensity of POWT upon addition of 0.5 and 1 equivalents of peptide shows that the affinity is high and in the low to medium nM range. The appearance of the emission spectrum of a mixture of JR2E and POWT is similar to that of POWT in an alkaline solution where the amino groups are unprotonated\(^{[52]}\). This indicates that the change in conformation of POWT is triggered by electrostatic interactions between negatively charged glutamic acids in JR2E and the positively charged amino groups in the side chains of the polymer. The decrease in intensity is due to aggregation of the polymer chains\(^{[52]}\), and the strong tendency to aggregate was striking in the preparation of a sample of JR2E and POWT for NMR spectroscopy. At the concentration required for the experiment, the JR2E-POWT complex precipitated almost immediately clearly showing that the interaction between JR2E and POWT forced the polymer chains to aggregate.
Fluorescence emission spectrum of 5.3 µM POWT (∗) and of POWT mixed with 0.5 equivalents of JR2K (○), 0.5 equivalents of JR2E (△), 0.25 equivalents of JR2E followed by 0.25 equivalents of JR2K (□), 0.25 equivalents of JR2K followed by 0.25 equivalents of JR2E (■), 1.0 equivalent of JR2K (♦), and 1.0 equivalent of JR2E (▲) after 10 min of incubation in a 20 mM phosphate buffer, pH 7.4.

The interactions between JR2E and POWT clearly affected the conformation of the polymer, but according to CD spectroscopic measurements the conformation of JR2E remained unordered, although it can not be excluded that alternative unordered conformations were populated in the complex. The CD signals of JR2E were almost identical whether the polymer was present or not (Figure 21). The nature of the interaction is complex. While it would not have been surprising to find that the peptide would adopt an ordered conformation in the complex due to the interaction with the chiral polymer, this was not observed. Also, the zwitterionic polymer could have shielded the negative charges at the dimer interface to enable helix-loop-helix dimer formation, in a ternary complex, but the interaction between JR2E and POWT leads to a polymer with negative net charge, apparently preventing JR2E from forming helical homodimers. The stoichiometry of the complex formed from peptide and POWT required further investigations (see chapter 6, section 6.2.).

The emission spectrum of POWT at pH 7.4 was also affected by JR2K (Figure 20). The appearance of the spectrum of JR2K and POWT was different from e.g. the
spectrum of JR2E and POWT. JR2K caused an increase in the intensity of the emission instead of a decrease, indicating that JR2K selects a different set of conformers of the polymer than JR2E. This behaviour of the polymer has been observed under conditions where the chains are separated from each other\cite{52}, indicating that the positively charged residues in JR2K forces the polymer chains to separate from each other rather than aggregate, which was the case with JR2E.

![Figure 21](image-url)

**Figure 21.** The CD spectrum of 5.3 µM of POWT (×) and the CD spectrum of POWT mixed with 1.0 equivalent of JR2E (△), 1.0 equivalent of JR2K (○), 0.5 equivalents of JR2E followed by an addition of 0.5 equivalents of JR2K (□) after 10 min of incubation in a 20 mM sodium phosphate buffer, pH 7.4. CD spectra of peptide solutions without POWT are labeled with filled symbols and do not deviate significantly from those recorded in the presence of POWT.

As in the case of JR2E, the polymer did not seem to induce a more ordered conformation in JR2K according to CD measurements (Figure 21). The CD spectrum of JR2K, too, was essentially the same whether the polymer was present or not. The reasons for the inability of JR2K to fold in the presence of the polymer were probably the same as those for JR2E. The interaction between the peptide and the polymer gives rise to a complex with a positive net charge, preventing JR2K from dimerising and folding, and in the peptide POWT complex, the peptide does not adopt a helical conformation.

Although the conformations of JR2E and JR2K were not affected by the interactions with POWT there was no doubt that complexes were formed. NMR measurements showed that JR2K was affected by the presence of the polymer. In the spectrum of
JR2K (Figure 22) the resonances are narrow and the shift dispersion poor, whereas in the NMR spectrum of JR2K together with POWT (Figure 22) the resonances are broadened due to complexation between the peptide and the polymer resulting in slower tumbling and shorter relaxation times.

The addition of JR2K to a solution of JR2E and POWT results in an increase in the intensity of the emission (Figure 20) indicating that the polymer interacts with the four-helix bundle and that a new complex containing all three components is formed.

![Figure 22. ¹H NMR spectra (600 MHz) of JR2K (upper left), JR2E/JR2K (upper right), JR2K/POWT (lower left) and JR2E/JR2K/POWT (lower right) in ²H₂O at pH 7.4 and 298 K.](image)

According to CD spectroscopy, a mixture of JR2E and JR2K is helical with and without POWT (Figure 21) although there is a slight difference in the ratio of the mean residue ellipticities at 208 nm and 222 nm. Apparently the interaction between the polymer and the four-helix bundle causes a slight distortion of the helical structure. NMR spectroscopic measurements show that there is an interaction between the four-helix bundle and the polymer, in analogy with experiments carried out with JR2K and POWT (Figure 22). The ability of POWT to discriminate between JR2E, JR2K and the heterodimer formed from JR2E and JR2K or alternatively, the ability of
JR2E, JR2K and the heterodimer formed from JR2E and JR2K to select different conformations of POWT was remarkable and suggested that this self-assembled complex could be used in biotechnical applications requiring high-affinity molecular recognition and sensitive detection.
6. JR2E and POWT as a platform for biosensor applications

A biosensor is an analytical device that utilises a biochemical mechanism to recognise and quantify a specific target molecule (analyte)\[^53\]. The recognition of the analyte may be mediated by different kinds of macromolecules e.g. enzymes\[^54\], antibodies\[^55, 56\], polymers\[^57\] and even microorganisms\[^58\]. A key feature of a biosensor is that the biochemical interaction with the analyte should induce a change in a chemical or physical parameter e.g. wavelength, frequency, refractive index or pH, that can be transformed to a measurable signal. A biosensor should be selective and able to discriminate between substrates and be sensitive and able to detect small quantities of the analyte. To meet this goal a polypeptide that is readily modified with high affinity ligands for various target molecules was combined with a conjugated polymer with optical properties that are sensitive to perturbations of the chemical environment due to amplification by a collective system response. To test the concept the well-characterised interaction between the enzyme human Carbonic Anhydrase II (HCAII) and its benzenesulphonamide inhibitor\[^59\] was used as a model system.

6.1. The transformation of JR2E into a binder of HCAII

Previous studies had shown that a designed polypeptide similar to JR2E could be transformed into a binder for HCAII upon modification with a benzenesulphonamide group attached via an alkyl linker\[^28\]. The length of the linker was found to highly influence the affinity between HCAII and the polypeptide and an optimal linker was found to contain between 5 to 7 methylene groups\[^28, 60\]. As a starting point in the design of a high-affinity binder for HCAII a linker with 5 methylene groups was chosen to attach a benzenesulphonamide residue to JR2E.

6.1.1. The synthesis of JRE2-lig

The peptide was synthesised according to a standard Fmoc chemistry protocol using acid labile protecting groups except for the lysine in position 14 that was protected by an allyloxycarbonyl group. This protection group can be selectively removed after the
synthesis by Pd[PPh₃]₄ while the peptide is still attached to the solid support (Scheme 2).

The ligand was prepared from 4-sulfamoyl-benzoic acid and 6-aminohexanoic acid using a carbodiimide mediated coupling reaction (Scheme 2). The ligand was attached to the deprotected lysine, and the peptide cleaved from the solid support by trifluoroacetic acid and purified by reversed phase HPLC.

Scheme 2. Synthesis of JR2E-lig. Reagents: (i) Dioxane/DMF 5:2 0 °C, DCC 5 °C; (ii) Acetone, Aqueous borate (pH 8.5); (iii) CHCl₃/HOAc/NMM 17:2:1, Pd[PPh₃]₄; (iv) DCM, DIPEA, EDC, HOBt.

6.2. Optimisation of the biosensor

The preferred function of a biosensor is that when no analyte is bound the emission should be weak or absent and when the analyte is captured the system lights up. Based on the observation that JR2E gave rise to low intensity and that the JR2E/JR2K dimer increased the intensity, JR2E was selected as the capture element of the biosensor.
with the expectation that when the functionalised form of JR2E would bind a target protein the complex would light up. The emission of POWT was expected to be quenched due to aggregation caused by the interaction with JR2E-lig. When the analyte HCAII was introduced the interaction between the functionalised peptide and the analyte was expected to cause a conformational change in the polymer leading to disruption of the aggregates and an increase in the intensity. For a system based on self-assembly it is important that the ratio of components, in this case polymer and peptide, is optimised and based on the correct stoichiometry. If the peptide concentration is too low compared to the polymer, the baseline will be high and there will be unbound polymer chains that can interact unspecifically with the analyte. On the other hand, if the peptide concentration is too high compared to the polymer then there will be an excess of peptide that binds the target protein but do not interact with the polymer. These peptides will interact with the analyte without giving rise to a change in emission and give rise to an insensitive system. To find an optimal ratio between the polymer and the peptide, POWT was titrated with JR2E-lig (Figure 23). From the appearance of the titration curve it is reasonable to believe that the polymer interacts with the peptide in a 2:1 ratio since the fluorescence signal is not further quenched after the concentration of JR2E-lig has reached a level that is approximately half of that of POWT.

![Figure 23](image)

**Figure 23.** Emission intensity of 100 nM POWT at 590 nm as a function of JR2E-lig concentration in 20 mM Trizma, pH 7.5.

### 6.3. The interactions between the peptide/polymer complex and HCAII

The titration of the POWT/JR2E-lig complex with HCAII showed that there is a distinct increase in the emission intensity upon increased HCAII concentration to a level where the concentration of HCAII is approximately equal to that of the affinity
ligand (Figure 24). NMR experiments with $^{15}$N-labeled HCAII in the presence of polypeptides modified with the same affinity ligand have shown previously that several residues around the binding cleft of HCAII are affected by the peptide scaffold when the ligand interacts with the binding pocket\textsuperscript{[60]}. These interactions probably alter the conformation of the peptide that in turn causes the disruption of the polymer/peptide aggregates to some extent, leading to an increased intensity of the emitted light. When the HCAII concentration exceeds the concentration of JR2E-lig all the affinity ligands are occupied and there is no further increase in the signal indicating that it is the specific interaction between the affinity ligand and HCAII that is responsible for the increased intensity (Figure 24).

A universal problem in biosensor applications is the occurrence of unspecific interactions between the target analyte and everything else in the sample but the reporter. This may cause false positive signals and unreliable data. In this system the interaction that is supposed to lead to a detectable signal is the interaction between HCAII and the polypeptide functionalised with the benzenesulphonamide and linker. It is possible that HCAII interacts with the polymer or polymer peptide complex, to cause a similar response without the mediation of the sulphonamide ligand. To determine if unspecific interactions contributed to the signal observed in the titration of POWT/JR2E-lig with HCAII the same experiment was carried out with POWT complexed to the unmodified peptide JR2E. The result showed that there are unspecific interactions but that they do not become significant until the HCAII concentration exceeds approximately twice the peptide concentration (Figure 24).

Figure 24. Emission intensities at 590 nm for POWT/JR2E-lig (100/50 nM) (●) and for POWT/JR2E (100/50 nM) (○) as a function of the concentration of HCAII in 20 mM Trizma, pH 7.5.
The addition of HCAII to a solution containing POWT also causes the polymer to aggregate, an effect that is observable as a decrease in the emission intensity (Figure 25). For this reason it reasonable to believe that the complex between the polymer and the polypeptide is not completely disrupted upon interaction with HCAII. In both cases when either POWT/JR2E-lig or POWT/JR2E was titrated with HCAII the emission intensity levelled out at high protein concentration (Figure 24). If the polymer/peptide complex dissociated upon interaction with HCAII, and the increased signal was due to the released polymer there would be a decrease in the signal at high protein concentration due to the interaction between the free polymer and HCAII. Since there are no signs of a decrease in the intensity of the emission at high protein concentration a complex appears to be formed containing all three components. The fact that the interaction between JR2E-lig and POWT is not completely broken upon interaction with the target analyte is an advantage since the otherwise somewhat promiscuous polymer is not free to interact with other substances in solution.

![Graph](image)

**Figure 25.** Fluorescence emission spectrum of 100 nM **POWT (∆)** 100 nM **POWT** and 100 nM **HCAII (○)** 100 nM **POWT** and 100 nM **JR2E-lig (◊)** in 20 mM Trizma, pH 7.5.

### 6.4. Screening for high-affinity protein ligands

If a ligand for a protein of interest is known, new ligands with higher affinity for the same binding site can be searched for in competition experiments provided the binding event can be monitored. This is of interest in the early stages of the development of new drugs where large substance libraries are screened for affinity against the target protein. The peptide JR2E functionalised with a known ligand for a protein target, complexed with POWT was expected to be a useful tool for high-affinity screening. Potential substances with higher affinity than the functionalised polypeptide would cause a significant decrease in the fluorescence intensity upon
competing successfully for the binding site since the release of the functionalised peptide would force the polymer to aggregate. In a proof-of-principle demonstration a mixture of POWT/JR2E-lig and HCAII was titrated with 2-acetylamido-1,3,4-thiadiazole-5-sulphonamide (acetazolamide), a high-affinity heterocyclic inhibitor of human Carbonic Anhydrase II\[61\]. In the absence of acetazolamide the system is lit up due to the interaction between the peptide and HCAII resulting in a twisted conformation of the polymer (Figure 26). Upon titration with acetazolamide the intensity decreased as JR2E-lig was released from the protein and the polymer aggregated (Figure 26). The result shows that the polymer/peptide complex behaves as expected and may be of use in screening applications.

**Figure 26.** Fluorescence emission intensities of a mixture of 50 nM HCAII 100 nM POWT and 50 nM JR2E-lig as a function of acetazolamide concentration in 20 mM Trizma, pH 7.5. The release of JR2E-lig from the complex with HCAII is observed as a decrease in intensity.
7. The use of protein-protein interactions in tailoring properties of hybrid materials

Colloidal particles of gold are promising building blocks for the development of functional materials and devices due to the possibility of chemically altering and controlling their function at the molecular level. Especially their optical properties make them interesting in e.g. biosensor applications\textsuperscript{[62]}. By replacing the valine in the loop sequence of JR2E with a cysteine (JR2EC) it was possible to covalently coat gold nanoparticles with the polypeptide. In a previous study it was shown that the folding properties of JR2EC in solution was the same as those of JR2E indicating that the cysteine in the loop does not affect the ability of JR2EC to form homo- and heterodimers\textsuperscript{[63]}. It was also shown that JR2EC was able to form homodimers on a planar gold surface at pH 4-5\textsuperscript{[63]}. As a control, a polypeptide without folding ability had been designed (JR2ECref) by replacing all L-Ala residues in JR2EC with D-Ala residues (Figure 27 inset).

![Figure 27](image)

**Figure 27.** Absorption versus pH of JR2EC (-) and JR2ECref (---) attached to gold nanoparticles. The lines are drawn as a guide for the eye only. (Inset) Mean residue ellipticity at 222 nm versus pH for 100 \( \mu \)M JR2EC (●) and JR2ECref (■) in aqueous solution.
Polypeptide functionalised nanoparticles dispersed in buffer solution showed a strong absorption with a maximum at approximately 520 nm in the pH range between pH 5 and pH 7 (Figure 27). In this pH range JR2EC and JR2ECref were negatively charged and repulsive forces prevented the particles from aggregating. At a pH below 5 the charged glutamic acids are protonated and a distinct red shift was observed for both JR2EC- and JR2ECref-functionalised gold particles due to aggregation (Figure 27). However, for JR2ECref-functionalised nanoparticles a much larger red shift was observed indicating a smaller interparticle distance \[^{[64]}\]. Since JR2ECref is unable to fold it was expected to collapse in an unordered conformation on the surface of the particle and allow the nanoparticles to come close to one another once JR2ECref was neutralised. JR2EC on the other hand is capable of forming homodimers that are rigid enough to keep the particles separated at a somewhat larger distance, and hence to give rise to a smaller red shift. Below pH 3.5 most of the glutamic acid residues are protonated and both polypeptides adopted a positive net charge. Charge-charge repulsion disrupted the aggregates resulting in a blue shift of the absorption maximum to 520 nm (Figure 27).

These results show that it is possible to tailor the optical properties of functionalised nanoparticles by using protein-protein interactions and to use synthetic polypeptides to control the assembly of nanoparticles.
8. Methods

8.1. Solid-phase peptide synthesis

In solid-phase peptide synthesis (SPPS), first introduced in 1963 by R. B. Merrifield\cite{65}, the first amino acid is covalently linked by its C-terminal end to an insoluble polymeric support. Side chain protected amino acids are coupled one by one to the growing polypeptide chain by the use of coupling reagents (Figure 28). A major advantage of SPPS is that after each coupling step excess amino acids and coupling reagents are conveniently removed. Other advantages are that branched polypeptide chains can be synthesised and unnatural amino acids can be introduced. SPPS is divided into two main strategies depending on which protecting group that is used to protect the $\alpha$-amino functional group of the amino acid. In Boc chemistry the tert-butyloxycarbonyl (Boc) group is used for N-$\alpha$ protection. This protecting group is removed by the use of strong acid such as trifluoroacetic acid (TFA). After the synthesis the peptide is cleaved from the solid support by hydrogen fluoride (HF). In the other strategy, which does not require such harsh conditions, 9-fluorenylmethyloxycarbonyl (Fmoc) is used for N-$\alpha$ protection. This protecting group is removed under mild basic conditions by the use of 20 % piperidine in N, N-dimethylformamide (DMF) and after the synthesis the polypeptide is cleaved from the solid support by TFA. All polypeptides described in this thesis are synthesised according to Fmoc chemistry protocols.

8.2. High performance liquid chromatography (HPLC)

HPLC is a convenient method for polypeptide purification and is based upon interactions between a solid stationary phase and a liquid mobile phase. The stationary phase can either be polar (straight phase) or unpolar (reversed phase). The most common stationary phase used for peptide purification is the reversed phase and the mobile phase is usually an aqueous solution containing methanol, isopropanol or acetonitrile. The polypeptides are usually detected by UV spectroscopy at 210-230 nm due to the absorption of the backbone amide bonds. In this work the polypeptides were purified on a semi preparative C8 column and eluted with mixtures of water and isopropanol containing 0.1 % TFA (v/v) and detected at 229 nm.
Figure 28. Simplified reaction scheme for solid-phase peptide synthesis using Fmoc chemistry with the first amino acid attached to the solid support. $P_1$, $P_2$ = acid-labile side-chain protecting groups.

8.3. Matrix assisted laser desorption ionisation time of flight (MALDI-TOF)

MALDI-TOF is a powerful and easy method for determining the molecular weight of large non-volatile molecules such as polypeptides and proteins\cite{66, 67}. The analyte is co-crystallised with a vast excess of a matrix compound, in this work $\alpha$-cyano-4-hydroxycinnamic acid, on a metal plate and introduced into a high vacuum chamber. The sample is then irradiated with nanosecond laser pulses causing the sample to
evaporate. In the process the matrix molecules are ionised and the analyte molecules are ionised by gas phase proton transfer. About 100-500 ns after the laser pulse a strong acceleration field is applied, bringing the same kinetic energy to the ions. The ions are travelling down a flight tube where they hit a detector and since the ions have the same kinetic energy the lighter molecules will arrive first. The mass-to-charge ratio is related to the time it takes an ion to reach the detector by

\[
\frac{m}{z} = \frac{2t^2 K}{L^2}
\]

where \(m\) is the mass, \(z\) is the number of charges, \(t\) is the flight time, \(K\) is the kinetic energy and \(L\) is the length of the flight tube.

MALDI-TOF was used in this work to identify all the synthesised polypeptides.

### 8.4. Circular dichroism (CD) spectroscopy

Plane polarised light can be thought of as being made up of two vectorial components of equal magnitude, one rotating clockwise (right handed, R) and one rotating counterclockwise (left handed, L). If plane polarised light is used to irradiate an optically active sample at a wavelength where it is absorbed the resulting radiation is elliptically polarised since there are different extinction coefficients for the right handed and the left handed circularly polarised light. The phenomenon is called circular dichroism\(^{[25]}\). In proteins the main contributors to the CD spectrum are the amide groups of the polypeptide backbone, which absorb light below 240 nm. Depending on their relative orientation in space in structural elements such as \(\alpha\)-helices and \(\beta\)-sheets they will give rise to characteristic CD spectra for folded proteins. For example the \(\alpha\)-helix gives rise to three characteristic signals in the CD spectrum one maximum at 192 nm and two minima at 208 and 222 nm.

As mentioned above, the CD signal arises from the differences in the extinction coefficients for the right handed and the left handed circularly polarised light, \(\Delta\varepsilon\), which is proportional to the difference in absorbance, \(\Delta A\), according to Lambert-Beers law;

\[
\Delta A = l \cdot c \cdot \Delta\varepsilon
\]
where \( l \) is the path length (cm) and \( c \) is the sample concentration (g/mL). The output is often converted to ellipticity, \( \Theta \), measured in degrees and related to \( \Delta A \) according to

\[
\Theta = 33 \cdot \Delta A
\]

Another term that is frequently used is the mean residue ellipticity \([\Theta]\) that is defined as

\[
[\Theta] = \frac{\Theta \cdot mrw}{10 \cdot l \cdot c}
\]

where \( \Theta \) is the observed ellipticity and \( mrw \) is the mean residue weight (g/mol).

CD spectroscopy was used to estimate the conformations of the different polypeptides in different chemical environments.

### 8.5. Nuclear magnetic resonance (NMR) spectroscopy

NMR spectroscopy is a technique that has many applications in the study of proteins e.g. in structural determinations or in the study of dynamic processes involved in ligand binding. The technique makes use of the fact that some nuclei have spin quantum numbers that are not equal to zero. The proton, the lightest isotope of hydrogen, represented by the symbol \(^1H\), has spin quantum numbers of \(+1/2\) and \(-1/2\). In the absence of an external magnetic field the associated magnetic moments are randomly distributed. However, if a magnetic field is applied the spins align either parallel or antiparallel to the magnetic field since the energy levels are split due to quantisation. The difference between the populations of the two energy levels is given by the Boltzmann distribution. If a radio frequency pulse is applied orthogonally to the primary magnetic field the Boltzmann distribution will be perturbed and the relaxation to equilibrium, the Free Induction Decay (FID) contains the spectral information. Fourier transformation of the FID converts it to a frequency spectrum where the resonances show up as peaks at chemical shifts determined by the chemical environments of the nuclei. Through-bond connectivities between nuclei give rise to fine structure that provides information about the structure of the molecule. In proteins there are many nuclei and overlapping signals that makes it difficult to assign a chemical shift to each atom in a simple spectrum. To overcome this problem multi-dimensional NMR techniques have been developed that separate the signals in two or
more dimensions to increase the resolution \cite{68}. However, simple one-dimensional $^1$H NMR spectra can provide important qualitative information about structure and dynamics of polypeptides and proteins. In a random coil for example the inversion rate is fast on the NMR time scale\cite{31} and the resulting chemical shifts are the weighted mean values of the chemical shifts of the interconverting conformers resulting in sharp lines with poor shift dispersion. In a more ordered conformation the interconversion rates are decreased resulting in broadened lines with increased shift dispersion. In a native protein the well-defined structure results in sharp lines and well dispersed chemical shifts. The inspection of a one-dimensional $^1$H NMR spectrum, recorded in a few minutes thus makes it possible to discriminate between conformational states of the polypeptide chain.

In this work one-dimensional $^1$H NMR spectra were used to obtain qualitative information about protein-protein interactions and protein-polymer interactions.

### 8.6. Fluorescence spectroscopy

Fluorescence spectroscopy is a technique that is widely used for studying conformational changes and protein interactions in biological systems. Fluorescence spectroscopy has become an important tool in the study of proteins because it is a very sensitive technique and because there is a vast number of different fluorophores available for the labelling of molecules and cells. Fluorescence occurs when a fluorophore absorbs light and is excited to a higher electronic state. The excited molecule rapidly loses some of the absorbed energy upon rotational and vibrational relaxation of the excited state. This process is nonradiative i.e. no light is emitted and instead heat is released. Once the molecule has reached the lowest energy level of the excited state a photon is emitted, which is detected as fluorescence. Since there is a nonradiative relaxation in the excited state the emitted light is always of longer wavelength than the absorbed light. The difference in wavelength between the emitted and absorbed light is referred to as the Stokes’ shift.

Fluorescence spectroscopy was used in this work to measure the conformational change of an organic polymer upon interaction with designed polypeptides.

### 8.7. Analytical ultracentrifugation

Analytical ultracentrifugation can be used to study proteins under various conditions and important information about molecular mass, oligomeric state and dissociation constants can be obtained\cite{27, 69}. There are mainly two different methods for studying
proteins with analytical ultracentrifugation, equilibrium sedimentation and velocity sedimentation. In this work only the equilibrium sedimentation method was used and a brief explanation of the method will be given here.

When a sample is centrifuged the sample molecules are forced away from the centre of rotation by the centrifugal force. At the same time there is a diffusion force, working against the centrifugal force, trying to establish a uniform concentration in the sample compartment. After some time the two opposing forces reaches an equilibrium leading to a gradient of increasing concentration away from the centre of rotation. By monitoring this gradient at equilibrium at different rotor speeds the parameters mentioned above can be determined.
9. References

27. Ralston, G. *Introduction to Analytical Ultracentrifugation*, Beckman Instruments Inc., Fullerton, **1993**.


10. Acknowledgments

Jaha då börjar den här resan gå mot sitt slut och det är många man vill tacka så för att inte glömma någon vill jag först och främst rikta ett stort tack till alla som på ett eller annat sätt bidragit till att denna bok blivit till.

Ett stort tack till…

Lars Baltzer, min handledare, för att du med din stora entusiasm och uppmuntran alltid lyckas se ljuspunkter även när det ser mörkt ut och för att du alltid tar dig tid att reda ut mer eller mindre intelligenta frågeställningar.

Gunnar Dolphin för alla intressanta diskussioner rörande peptidsyntes.

Martin Lundqvist för att du alltid tog dig tid att hjälpa mig med mina NMR-försök samt fixa till CD-maskinen när den bråkade.

Peter Nilsson för ett intressant och trevligt samarbete i samband med peptid/polymerarbetet.

Sofia Håkansson Hederos för att du alltid bryr dig och för alla trevliga lunchdiskussioner och fikastunder.

Susanne Andersson som håller skeppet flytande och alltid hjälper till när krångliga papper ska fyllas i.


Lärare, doktorander och teknisk och administrativ personal på kemiavdelningen i Linköping för att ni skapar en trivsam atmosfär att arbeta i.
Lotta och Johan i Uppsala för att man alltid känner sig välkommen när man kommer och hälsar på.

Jag vill jag även tacka mina föräldrar, Iris och Lars Rydberg, och min syster, Ann-Marie Rydberg, som alltid stöttar och ställer upp i alla väder.

Slutligen vill jag tacka min älskade fru Viktoria Rydberg som alltid tror på mig även när jag själv tvivlar.