

# **Raman Spectroscopy of Proteins, DNA and Their Complexes**

Inaugural Dissertation  
zur Erlangung des akademischen Grades

“Doktor der Naturwissenschaften”

- Dr. rer. nat. -

am Fachbereich Biologie, Chemie, Pharmazie  
der Freien Universität Berlin

Vorgelegt von  
Lubomír Dostál

**Berlin 2004**

Diese Arbeit wurde im Zeitraum von September 2001 bis November 2004 in der Arbeitsgruppe Biopolymerspektroskopie (Leiter: Prof. Dr. Heinz Welfle) am Max-Delbrück-Centrum für Molekulare Medizin Berlin-Buch angefertigt.

**1. Gutachter: Prof. Dr. Heinz Welfle**

**2. Gutachter: Prof. Dr. Udo Heinemann**

**Eingereicht am: 10. November 2004**

**Tag der mündlichen Prüfung am: 21. Januar 2005**

# Table of Contents

Abstract .....	1
Zusammenfassung .....	4
Acknowledgments .....	7
Abbreviations .....	8
1. Introduction .....	10
1.1 Bacterial plasmids .....	11
1.1.1. Plasmid Maintenance .....	11
1.1.2. Plasmid Partitioning .....	11
1.1.3. Postsegregational Killing Mechanism .....	12
1.1.4. <i>inc18</i> Family and Plasmid pSM19035 .....	13
1.2. $\omega$ Protein .....	14
1.2.1. Properties and Crystal Structure .....	14
1.2.2. Ribbon-Helix-Helix Binding Motif .....	15
1.2.3. Binding Site and Properties of $\omega_2$ -Operator DNA Complex .....	17
1.3. Arc Repressor .....	18
1.4. Regulatory Protein KorB .....	20
1.5. DNA-Binding Protein Sac7d .....	21
2. Objectives .....	24
3. Materials and Methods .....	26
3.1. Bacterial Strains and Plasmids .....	26
3.2. LB Medium for <i>E. coli</i> .....	26
3.3. FPLC columns and devices .....	26
3.4. Buffers .....	27
3.5. Determination of Protein Purity .....	27
3.6. Expression and Purification of $\omega$ and Arc Proteins .....	27
3.7. Oligonucleotide Preparation .....	29
3.8. Sample Preparation .....	30
3.8.1. $\omega$ and Arc-DNA Complexes .....	30
3.8.2. KorB-operator DNA complexes .....	31
3.8.3. Sac7d-DNA Complex .....	31
3.8.4. Hydrogen/Deuterium Exchange .....	32

3.9. Determination of Complex Formation.....	32
4. Raman Spectroscopy.....	34
4.1. Raman Effect.....	34
4.2. Raman Spectrometer .....	36
4.3. Spectra Measurements .....	38
4.3.1. Raman spectra .....	38
4.3.2. Circular Dichroism (CD), Fluorescence, and Concentration Measurements....	40
4.4 Interpretation of Raman Spectra .....	40
4.4.1. Evaluation of Raman Spectra.....	41
4.4.2. Raman Spectra of Proteins.....	41
4.4.3. Raman Spectra of Nucleic Acids .....	45
4.4.4. Raman Spectra of Protein–DNA Complexes.....	47
5. Results and Discussion.....	48
5.1. Results and Interpretation: Omega Protein .....	49
5.1.1. Raman Signature of Oligonucleotides .....	49
5.1.2. Raman Signature of Omega Protein.....	54
5.1.3. Raman Signatures of the $\omega_2$ -DNA Complexes .....	58
5.2. Discussion: Omega Protein.....	73
5.3. Results and Interpretation: Arc Repressor .....	77
5.3.1. Fluorescence and CD Measurements of Arc and Arc–DNA Complexes .....	77
5.3.2. Raman Signature of Arc-wt and Arc-F10H Repressor Proteins .....	78
5.3.3. Raman Signature of the Arc Repressor Operator DNA .....	81
5.3.4. Raman Analysis of the Arc Repressor–Operator DNA complex.....	82
5.3.5. The Arc-wt in Complex with Non-Operator DNA .....	86
5.3.6. The Interaction of Arc-F10H with Non-Operator DNA .....	88
5.3.7. Time Dependence of Arc-F10H Binding to Non-Operator DNA.....	91
5.4. Discussion: Arc Repressor .....	95
5.4.1. Arc-Operator DNA Complex.....	95
5.4.2. Arc–Non-Operator DNA complex.....	97
5.5. Results and Discussion: Regulatory Protein KorB .....	99
5.5.1 Raman Spectrum of the Unbound 17-bp $O_B$ DNA .....	99
5.5.2. Raman Spectrum of KorB-O.....	100
5.5.3. Raman Analysis of the KorB– and KorB-O–DNA complexes.....	103
5.6. Results and Interpretation: Sac7d Protein .....	110

5.6.1. Raman Spectrum of d(GAGGCGCCTC) <sub>2</sub> .....	110
5.6.2. Raman Signature of Sac7d .....	111
5.6.3. Raman Analysis of the Sac7d–d(GAGGCGCCTC) <sub>2</sub> Complex .....	116
5.7. Discussion: Sac7d Protein .....	121
6. References .....	124
Curriculum Vitae .....	132
List of Publications .....	133
Selbständigkeitserklärung .....	134

## Abstract

Protein–DNA recognition plays an extremely important role for the genetic regulation in all living organisms. Proteins are involved in the control of DNA replication, transcription, repair, condensation, packaging, and transport. A complete understanding of the biological functions of protein–DNA interaction requires knowledge of dynamic and static properties of protein–DNA complexes, such as three dimensional structures, the relevance of the structure to environmental conditions, kinetics, thermodynamics, and many others. Among the multitude of methods that are necessary for the achievement of further progress, Raman spectroscopy is a valuable tool with specific merits and contributions in the field. We applied Raman spectroscopy for the analysis of  $\omega_2$ –operator DNA complex, Arc–operator DNA complex, KorB–operator DNA complex, and Sac7d–decamer d(GAGGCGCCTC)<sub>2</sub> complex.

The pSM19035-encoded dimeric  $\omega$  protein ( $\omega_2$ ) regulates the transcription of genes required for the control of plasmid copy number and stable inheritance. It is a member of the MetJ/Arc structural superfamily of repressors featuring a ribbon-helix-helix DNA binding motif.  $\omega_2$  binds specifically to promoter regions containing at least two consecutive copies of a 5'-<sup>A</sup>/TATCAC<sup>A</sup>/T-3' (top strand) sequence composed of seven base pairs. Binding of  $\omega_2$  to oligonucleotides composed of heptads is connected with significant changes in the Raman difference spectra (spectrum of complex minus spectra of DNA and  $\omega_2$  protein). The Raman markers indicate direct contacts with adenine, guanine, and thymine in the major groove of operator DNA suggesting that the central 5'-ATCAC-3' stretch of the heptads might be the main target site for  $\omega_2$  binding to operator DNA. The Raman data also indicate that Thr29 is essential for operator DNA binding and probably makes a hydrogen bond with one of the bases. Surprisingly, the spectral features are rather similar that were observed upon binding of  $\omega_2$  to oligonucleotides composed of different numbers (one, two or four) and different orientation (direct or inverted) of heptads. The results point to a high flexibility and adaptability of both  $\omega_2$  and DNA upon complex formation.

Solution properties of Arc repressors (wild type and F10H variant) from *Salmonella typhimurium* bacteriophage P22 and their complexes with operator and non-operator DNA were characterized by circular dichroism, fluorescence, and Raman difference spectroscopy. Raman markers in the difference spectra (spectrum of the complex minus

spectra of DNA and Arc) indicate binding of Arc in the major groove, several direct contacts, e.g. hydrogen bonds of protein side chains with bases, and slight perturbations of the deoxyribose ring systems that are consistent with bending of the operator DNA. Trp14, the only one tryptophan of Arc repressor monomers, serves as a very sensitive tool for monitoring properties of the hydrophobic core of the protein. The Raman spectra identify a largely different  $\chi^{2,1}$  rotation angle of Trp14 in the a Arc variant with His instead of Phe at position 10 (Arc-F10H) compared to that in wild type Arc. In complexes of Arc-wt and Arc-F10H with DNA, however, the Trp14  $\chi^{2,1}$  rotation angles are similar in both proteins. Furthermore, in both complexes a strengthening of the van der Waals interactions of the phenyl ring of Trp14 is indicated compared to these interactions in the free proteins. The Raman difference spectra of Arc repressor–non-operator DNA complex indicate unspecific interactions which are very different for both Arc-F10H and Arc-wt proteins.

KorB protein encoded by the conjugative plasmid RP4 is a global regulator controlling the expression of plasmid genes the products of which are involved in replication, transfer and stable inheritance. KorB is a homodimeric protein which binds to palindromic 13-bp DNA sequences (5'-TTTAGC(<sup>G</sup>/<sub>C</sub>)GCTAAA-3') present 12 times on the 60-kb plasmid. Details of the KorB– and KorB-O–operator DNA complex were investigated and compared with known crystal structure of KorB-O–operator DNA complex. The KorB-O subunit recognizes and binds to the operator sequence (O<sub>B</sub>). Comparison of the Raman spectra of free KorB-O domain and O<sub>B</sub> DNA with the spectrum of the complex reveals large differences. KorB-O binds in the major groove of the O<sub>B</sub> DNA and Raman markers clearly indicate hydrogen bonds to guanine N7 and O6. The KorB-O–DNA interactions induce changes in the DNA backbone and in the secondary structure of the protein. Wild type KorB and KorB-O bind in the same way to operator DNA as shown by the similarity of the Raman difference spectra of the complexes.

Members of the Sso7/Sac7 protein family and other related proteins are believed to play an important role for DNA packaging and maintenance in archeons. Sac7d is a small, abundant, basic and non-specific DNA-binding protein of the hyperthermophilic archeon *Sulfolobus acidocaldarius*. Structures of several complexes of Sso7d/Sac7d with DNA octamers have shown sequence unspecific minor groove binding of the proteins and sharp kinking of the double helix. Corresponding Raman vibrational signatures have been identified in this study. A Raman spectroscopic analysis of Sac7d binding to the oligonucleotide decamer d(GAGGCGCCTC)<sub>2</sub> reveals large conformational perturbations of the DNA structure upon complex formation. Large changes in the DNA backbone and

partial B- to A-form DNA transitions are indicated that are closely associated with a C2'-endo/anti to C3'-endo/anti conversion of the deoxyadenosyl moiety upon Sac7d binding. The major spectral feature of Sac7d binding indicates kinking of the DNA. Minor groove binding is displayed by guanine N2 and guanine N3 signals and supported by Trp24 features. Trp24 is the only tryptophan residue present in Sac7d. Binding of Trp24 to guanine N3 was demonstrated in crystal structures of Sac7d–DNA complexes. No changes of the Sac7d secondary structure have been detected upon DNA binding.



## Zusammenfassung

Protein-DNA-Erkennung spielt eine äußerst wichtige Rolle bei der Genregulation in allen lebenden Organismen, wobei die Proteine in die Kontrolle der DNA-Replikation, der Transkription, der Reparatur, der Kondensation, der Verpackung und des Transports eingebunden sind. Ein vollständiges Verständnis der biologischen Funktionen der Protein-DNA-Wechselwirkung erfordert die Kenntnis der dynamischen und statischen Eigenschaften von Protein-DNA-Komplexen, wie deren dreidimensionale Strukturen, die Relevanz der Strukturen zu den Umgebungsbedingungen, sowie kinetische, thermodynamische und andere Daten. Aus der Vielzahl von Methoden, die für den weiteren Erkenntnisfortschritt erforderlich sind, ist die Raman-Spektroskopie eine nützliche Methode von hohem Wert. Sie kann spezifische Beiträge zur Untersuchung von Proteinen, Nukleinsäuren und Protein-Nukleinsäure-Komplexen erbringen und wurde in der vorgelegten Dissertation für die Analyse von  $\omega_2$ -Operator-DNA-Komplexen, Arc-Operator-DNA-Komplexen, KorB-Operator-DNA-Komplexen und von Sac7d-Decamer d(GAGGCGCCTC)<sub>2</sub>-Komplexen eingesetzt.

Das pSM10035-kodierte dimere  $\omega$ -Protein ( $\omega_2$ ) reguliert die Transkription von Genen, deren Produkte für die Kontrolle der Anzahl an Plasmidkopien und für die stabile Vererbung der Plasmide benötigt werden.  $\omega_2$  gehört zur Superfamilie der MetJ/Arc-Repressoren, deren Mitglieder mit einem sog. „Ribbon-Helix-Helix“-Motif zur DNA-Bindung ausgestattet sind.  $\omega_2$  bindet mit hoher Spezifität an DNA-Heptaden der Sequenz 5'-<sup>A</sup>/TATCAC<sup>A</sup>/T-3' (Topstrang), die sich in der Promotorregion mehrfach wiederholen. Die Bindung von  $\omega_2$  an Oligonukleotide, die als Modelle für Operator-DNA mindestens zwei Heptaden enthalten, ist mit signifikanten Änderungen in den Raman-Spektren verbunden. Diese Änderungen werden in Differenzspektren deutlich, die durch Subtraktion der Spektren von DNA und  $\omega_2$  vom Spektrum des Komplexes berechnet werden. In den Differenzspektren sind spezifische Raman-Markerbanden zu sehen, die direkte Kontakte des Proteins mit Adenin, Guanin und Thymin in der großen Furche der Operator-DNA anzeigen. Diese Befunde sprechen für die Bedeutung des zentralen Basenabschnitts 5'-ATCAC-3' der Heptaden als entscheidendem Bindungsbereich für die Bindung von  $\omega_2$  an die Operator-DNA. Weiterhin ist aus den Raman-Daten abzuleiten, dass Thr29 essentiell für die Operator-DNA-Bindung ist und wahrscheinlich eine Wasserstoffbrücke mit einer der Basen bildet. Überraschenderweise sind die spektralen Eigenschaften der Komplexe

von  $\omega_2$  mit Oligonukleotiden unterschiedlicher Heptadenanzahl (eine, zwei bzw. vier) sowie unterschiedlicher Orientierung der Heptaden (direkt oder invers) weitgehend ähnlich, diese Befunde weisen auf eine hohe Flexibilität und Anpassung der Konformationen von  $\omega_2$  und DNA bei der Komplexbildung hin.

Der vom Bakteriophagen P22 aus *Salmonella typhimurium* kodierte Arc-Repressor und eine Variante von Arc (Arc-F10H) sowie deren Komplexe mit DNA wurden mittels Zirkulardichroismus, Fluoreszenz- und Raman-Differenzspektroskopie charakterisiert, um Eigenschaften der Proteine und der DNA-Komplexe in Lösung zu ermitteln. Es wurde ein spezifisch bindendes Oligonucleotid, dessen Sequenz von der des Arc-Operators abgeleitet war, sowie ein Oligonucleotid mit gleicher Basenzusammensetzung, aber zufälliger Sequenz eingesetzt. Raman-Marker in den Differenzspektren (Spektrum des Komplexes abzüglich der Spektren von DNA und Arc-Repressor) weisen auf die Bindung von Arc in der großen Furche, auf mehrere direkte Kontakte, z.B. Wasserstoffbrückenbindungen der Proteinseitenketten mit den Basen und auf geringe Veränderungen des Deoxyriboeringsystems hin und stehen in Übereinstimmung mit einer Krümmung der Operator-DNA im Komplex. Der Arc-Repressor enthält nur einen Tryptophanrest pro Monomer in Position 14 (Trp14), der als sehr empfindliche Sonde für die Charakterisierung der Eigenschaften des hydrophoben Kerns des Proteins eingesetzt werden kann. Aus den Raman-Spektren ergeben sich deutlich unterschiedliche  $\chi^{2,1}$  Rotationswinkel für Trp14 in der Arc-F10H-Variante im Vergleich zum Arc-Wildtyp. In den DNA-Komplexen der beiden Proteine sind jedoch die Tryptophane ähnlich orientiert. Verglichen mit den freien Proteinen wird in den Komplexen eine Zunahme der van der Waals-Wechselwirkungen des Trp14-Phenylringes beobachtet. Bei den hohen Konzentrationen im mM-Bereich, die für die Raman-Messungen eingesetzt wurden, binden Arc-Wildtyp und Arc-F10H auch an das Oligonucleotid mit zufälliger Sequenz. Die Raman-Differenzspektren dieser Komplexe zeigen unspezifische Wechselwirkungen an, die sich deutlich von den spezifischen Wechselwirkungen der Proteine mit Operator-DNA unterscheiden. Ebenfalls deutlich verschieden ist das Verhalten der von Arc-F10H und Arc-Wildtyp ausgebildeten Komplexe.

Das vom konjugativen Plasmid RP4 kodierte Protein KorB ist ein globaler Regulator, der die Expression von Plasmidgenen kontrolliert, deren Produkte an der Replikation, am Transfer und an der stabilen Vererbung des Plasmids RP4 beteiligt sind. KorB ist ein homodimeres Protein, das an eine 13 Basenpaare lange palindromische DNA-Sequenz (5'-

TTTAGC(<sup>G/C</sup>)GCTAAA-3') bindet, die zwölfmal im 60-kb-Plasmid vorkommt. Komplexe aus KorB und KorB-O mit Operator-DNA wurden Raman-spektroskopisch untersucht und mit der bekannten Kristallstruktur des Komplexes aus KorB-O und Operator-DNA verglichen. KorB-O ist eine rekombinant hergestellte Domäne von KorB, die die Operatorsequenz (O<sub>B</sub>) erkennt und bindet. Der Vergleich der Raman-Spektren der freien KorB-O-Domäne und der O<sub>B</sub>-DNA mit dem Spektrum des Komplexes läßt große Differenzen erkennen. KorB-O bindet in der großen Furche der DNA. Die Raman-Marker zeigen Wasserstoffbrücken zwischen KorB-O und Guanin-N7 und -O6 an. Die Wechselwirkungen von KorB mit der DNA induzieren Änderungen im DNA-Rückgrat und in der Sekundärstruktur des Proteins. Die Ähnlichkeit der Raman-Differenzspektren der Komplexe von Wildtyp-KorB und KorB-O mit Operator-DNA zeigt, dass beide Proteine in gleicher Weise an die Operator-DNA binden.

Es wird vermutet, dass Proteine der Sso7/Sac7-Familie und andere verwandte Proteine in Archebakterien eine wichtige Rolle bei der DNA-Verpackung und bei deren Aufrechterhaltung spielen. Sac7d ist ein kleines, weit verbreitetes, basisches und unspezifisch an DNA bindendes Protein aus dem hyperthermophilen Archebakterium *Sulfolobus acidocaldarius*. Kristallstrukturen verschiedener Komplexe von Sso7d/Sac7d mit DNA-Oktameren zeigen eine Bindung der Proteine in der kleinen Furche und einen scharfen Knick der Doppelhelix an. Entsprechende Raman-Banden wurden in der vorliegenden Arbeit identifiziert. Eine Raman-spektroskopische Analyse der Sac7d-Bindung an das Dekamer d(GAGGCGCCTC)<sub>2</sub> weist auf große Veränderungen der DNA-Konformation hin, die mit der Bindung des Proteins induziert werden. Es werden beträchtliche Änderungen im DNA-Rückgrat und partielle Übergänge von der B-DNA zur A-Form - nach Bindung von Sac7d beobachtet, die eng mit einem C2'-endo/anti zu C3'-endo/anti Übergang des Deoxyadenosylrestes verbunden sind. Die signifikanten spektralen Merkmale der Sac7d-Bindung deuten auf einen Knick in der DNA hin. Die Bindung in der kleinen Furche wird durch Guanin-N2 und -N3 Signale angezeigt und durch spektrale Eigenschaften des Trp24, dem einzigen Tryptophanrest in Sac7d, gestützt. Die Bindung von Trp24 an Guanin N3 wurde in Kristallstrukturen von Sac7d-DNA-Komplexen nachgewiesen. Es wurden keine Änderungen der Sac7d-Sekundärstruktur nach DNA-Bindung festgestellt.

## Acknowledgments

First, I would like to thank my supervisor Prof. Dr. Heinz Welfle for supporting this work with knowledge, ideas, criticism, and motivating discussions and excellent working conditions in AG Biopolymerspektroskopie in MDC.

Special thanks are to my second supervisor Prof. Dr. Udo Heinemann.

I'm grateful to all my colleagues in our lab for creating very friendly atmosphere and helping with the problems. Especially I very much appreciate the help of Dr. Rolf Misselwitz. He is a great teacher who taught me basics of protein biochemistry and molecular biology and helped me out with the german Zusammenfassung. I thank Brunhilde Kannen for guidance in the labs and technical assistance and Stefan Lättig for guiding my first steps in the Raman lab.

My sincere thanks go to all my collaborators Juan C. Alonso, Dheeraj Khare, Jiří Bok, Udo Heinemann, Erich Lanka, Chin-Yu Chen, and Andrew H.-J. Wang. Dheeraj Khare, Udo Heinemann and Erich Lanka were great help in preparing the KorB article. Jiří Bok programmed the calibration procedure. Chin-Yu Chen and Andrew H.-J. Wang assisted with the work on Sac7d. Juan C. Alonso helped with the work on Omega protein.

I was fortunate to have nice friends in Prague like Vladimír Baumruk and Vladimír Kopecký who provided many valuable advices.

I would also like to thank the MDC, Graduate College, DFG and EU for providing the resources and funding for my research.

My greatest thanks are reserved for my parents and family.

## Abbreviations

Å	Angstrom, $1 \text{ Å} = 10^{-10} \text{ m}$
A	adenine
Arc-F10H	Arc repressor variant with His instead of Phe in position 10
Arc-F10V	Arc repressor variant with Val instead of Phe in position 10
Arc-F10Y	Arc repressor variant with Tyr instead of Phe in position 10
Arc-wt	wild type Arc repressor;
bp	base pair
<sup>Br</sup> U	deoxybromouridine
C	cytosine
CCD	charge-coupled-device
CD	circular dichroism
Da	Dalton
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
FPLC	fast performance liquid chromatography
G	guanine
h	Planck's constant
HTH	helix-turn-helix
IPTG	isopropyl-β-D-thiogalactopyranoside
IR	infrared
KorB-N	N-terminal domain of KorB
KorB-O	DNA (operator) binding domain of KorB
mRNA	messenger RNA
MUT	DNA sequence (Figure 5.1)
NMR	nuclear magnetic resonance
O <sub>B</sub>	operator DNA
OD	optical density
PAGE	polyacrylamide gel electrophoresis
Par	partitioning
PDB	protein data bank

pI	isoelectric point
PSK	postsegregational killing
REF1	DNA sequence (Figure 5.1)
REF2	DNA sequence (Figure 5.1)
RHH	ribbon-helix-helix
RNA	ribonucleic acid
rpm	round per minute
SDS	sodium dodecyl sulfate
T	thymine
TA	toxin-antitoxin
Tris	Trishydroxymethylaminomethane
UV	ultra violet
$\alpha$	polarizability
$\mu$	dipole moment
$\nu$	frequency
$\omega_2$	Omega homodimer
$\omega_2$ -T29A	$\omega_2$ variant – threonine 29 substituted for alanine
$\omega_2$ - $\Delta$ N18	$\omega_2$ variant – first 18 N-terminal amino acids are deleted