

6. Zusammenfassung - Summary

6.1. Deutsch - German

6.1.1. Nicotinamid-Mononukleotid-Adenylyltransferase von *Homo sapiens* im Komplex mit Nicotinamid-Mononukleotid

Nicotinamid-Adenin-Dinukleotid (NAD⁺) ist ein bedeutendes Molekül als Coenzym bei zellulären Redox-Reaktionen und Signaltransduktionswegen. An seiner Biosynthese ist das Enzym Nicotinamid/Nicotinat-Mononukleotid-Adenylyltransferase (NMNAT/NaMNAT) beteiligt, indem es den Adenosin-Monophosphat-Teil von ATP auf Nicotinamid- oder Nicotinat-Mononukleotid (NMN/NaMN) überträgt. Im Fall von NaMN ist das Produkt Nicotinat-Adenin-Dinukleotid (NaAD⁺), welches von der NAD⁺-Synthetase zu NAD⁺ amidiert wird. Im Fall von NMN wird direkt NAD⁺ synthetisiert.

Die Kristallstruktur von NMNAT von *Homo sapiens* im Komplex mit NMN wurde mit einer maximalen Auflösung von 2.9 Å durch die SAD-Methode (single-wavelength anomalous dispersion) gelöst. Finale R_{cryst} bzw. R_{free}-Werte betragen 24.6 bzw. 28.6 %. Die Struktur von NMNAT besteht aus einem sechssträngigen parallelen β-Faltblatt mit Helices auf beiden Seiten, welches im Kern dem Rossmann-Faltungstyp entspricht. Elektronendichte wird weiterhin für den Liganden NMN beobachtet, nicht jedoch für eine Schleife von 37 Aminosäuren, Reste 109 bis 146, die strukturell ungeordnet sind. Von den homologen Proteinen aus *M. thermoautotrophicum* und *M. jannaschii* unterscheidet es sich durch diese Schleife, die eine nukleäre Lokalisationssequenz NLS enthält, und durch zusätzliche Aminosäuren, die im humanen Enzym die Helices **H** und **I** sowie den Strang **f** bilden. Diese Sekundärstrukturelemente verursachen ein unterschiedliches Oligomerisierungsverhalten als in den Proteinen aus Archaea, die jedoch alle als biologische Einheit ein Hexamer bilden. An Protein-Protein-Kontakten zwischen den Monomeren, die sich als zwei übereinander liegende Trimere formieren, sind die Helices **A**, **H** und **J** sowie die Schleife zwischen Helices **J** und **K** beteiligt. Die Trimere sind gegeneinander verdreht, die

Faltblätter von Untereinheiten verschiedener Trimere ordnen sich seitwärts gerichtet in einer antiparallelen Weise zu einer Art Super-Sekundärstruktur zusammen, ohne sich jedoch zu berühren. Bakterielle NMNAT kommt dagegen als Monomer (*E. coli*) oder Dimer (*B. subtilis*) vor und weist im Fall von *E. coli* NMNAT noch weitere zusätzliche Sekundärstrukturelemente auf.

An der Bindung des Liganden NMN sind die Aminosäuren Ser16, Lys57, Trp92, Thr95, Leu168 und Trp169 beteiligt, die sich alle auf einer Seite des β -Faltblattes befinden. Durch einen Vergleich mit weiteren, von anderen Gruppen publizierten Strukturen der humanen NMNAT (dem Apoenzym und dem NAD⁺-Komplex) sowie mit den Archaea- und Prokaria-Analoga wurde ein Model für die Ligandenbindung und Synthesereaktion entwickelt. Dabei befindet sich das Enzym in der Apo-Form in geöffnetem Zustand. Die Helices **F** und **G** auf der einen Seite und die Region um Helix **B** auf der anderen Seite sind weiter voneinander entfernt als im Komplex mit einem Liganden. In diesem Bereich der Ligandenbindungstasche, auf der einen Seite des Faltblattes, bindet NMN. ATP, der zweite Reaktionspartner, wird möglicherweise von einem positiv geladenen Cluster angezogen, der sich nur durch die Hexamer-Bildung formiert und könnte weitergeleitet werden bis zum Reaktionszentrum, eventuell unter Mitwirkung von Lys58 und Arg231. ATP bindet auf der anderen Seite des β -Faltblattes, beteiligt sind die Aminosäuren Phe17, Gly156, Glu215 und Asn219 sowie wahrscheinlich die konservierten Reste Thr21, His24, Ser222, Thr224 und Arg227, die wohl die β - und γ -Phosphate binden. Mit der Substratbindung nähern sich die oben genannten Elemente an und gleichzeitig ordnet sich der C-Terminus des Proteins (ab Rest 258). Er bedeckt dann Teile der NMN-Bindungstasche sowie die Aminosäuren His24 und Lys58. An der Synthesereaktion sind möglicherweise die Aminosäuren Gly15, Ser16, Phe17, His24 und Lys57 beteiligt, die in räumlicher Nähe der neu zu knüpfenden Phosphorester-Bindung liegen. Abdissoziation von NAD⁺ und Pyrophosphat PP_i gehen einher mit der Entspannung des Proteins und vervollständigen den Reaktionszyklus.

6.1.2. Domäne I der Homing Endonuclease PI-SceI von *Saccharomyces cerevisiae*

Die Homing Endonuclease PI-SceI von *S. cerevisiae* ist ein Intein, eine *interne Proteinsequenz*, eingebettet in die Extein-Sequenzen der 69 kDa-Untereinheit der vakuolären Membran-H⁺-ATPase. In einem Protein-Splicing-Prozess schneidet sie sich selbst aus dem Vorläuferprotein heraus und verbindet die beiden Exteine. Verantwortlich hierfür ist die Domäne I von PI-SceI, die gleichzeitig den Hauptteil der Bindungsenergie für eine spezifische DNA-Sequenz von mindestens 31 bp liefert. Die Erkennungssequenz ist der *VMA1-Dvde*-Locus des *S. cerevisiae* Chromosoms 8. Dieser Locus ist dasjenige Allel für die *VMA1*-Gensequenz (*VMA*: vacuolar membrane H⁺-ATPase), das defizient für *vde* ist (VDE: *VMA*-derived endonuclease = PI-SceI). Die Domäne II von PI-SceI besitzt das endonukleolytische Zentrum, welches die spezifische Erkennungssequenz schneidet und die Insertion des *vde*-Gens initiiert („homing“). Sie gehört zur LAGLIDADG-Familie von Homing Endonukleasen und bindet die DNA nur locker. Die Kristallstruktur der Domäne I von PI-SceI wurde mit der Methode des Molekularen Ersatzes bei einer maximalen Auflösung von 1.35 Å gelöst. Finale R_{cryst} bzw. R_{free}-Werte betragen 15.0 bzw. 18.9 %. Obwohl schon vier weitere Kristallstrukturen des gesamten Proteins bekannt waren, geben die hohe Auflösung und die gute Definition von Bereichen, die in den anderen Strukturen nicht oder schlecht interpretierbar waren, einen Einblick auf weitere Details. Die Ziele, die gesamte Proteinsequenz in der Elektronendichte zu beobachten und einen Protein-DNA-Kokristall zu erzeugen, wurden jedoch nicht erreicht. Vielmehr wuchsen in Komplex-Ansätzen lediglich Proteinkristalle, die umfangreiche Kristallkontakte ausbilden. 96 von 217 Aminosäuren sind daran beteiligt. Domäne I von PI-SceI, bzw. der Kernbereich, nimmt den Hint-Faltungstyp der Hedgehog/Intein-Domäne ein und besteht hauptsächlich aus β-Strängen. Das aktive Zentrum der Domäne I, die Protein-Splicing-Stelle, zeigt neben Cystein1 noch zwei N-terminale Extein-Reste, ihm fehlt jedoch mit Asn454 der Intein-C-Terminus. Das gängige Modell für den ersten Schritt des Protein-Splicing-Reaktionsmechanismus' kann strukturell bestätigt werden. Dabei initiiert die Seitenkette von Cys1 durch einen nukleophilen Angriff auf den Carbonyl-Kohlenstoff von Ala(-1) den initialen N-S Acyl-Shift, dem eine

Transesterifizierung, eine Succinimid-Bildung und der finale S-N Acyl-Shift folgen.

Die Reste Gln55 bis Glu66 wurden aufgrund lokaler Unordnung nicht in der Elektronendichte beobachtet. Sie bilden offenbar eine flexible Schleife, die im Verdacht steht, an der festen DNA-Bindung der Domäne I beteiligt zu sein. Die weiteren Reste, denen aufgrund biochemischer Daten eine Beteiligung an der Bindung zugerechnet wird, befinden sich alle an der Vorderseite der zangenartigen Subdomäne, die von den Strängen **i**, **j** und **k** sowie von den Helices **B** und **C** gebildet wird. Wahrscheinlich können auch die Helices **D** und **E** sowie der Strang **I** dazugezählt werden, wo ebenfalls potenziell DNA bindende Aminosäuren liegen. Eine strukturelle Überlagerung aller bekannten Proteinketten von PI-SceI Domäne I legt nahe, dass neben den flexiblen Resten Gln55 bis Glu66 auch die Aminosäuren Arg94, Arg90 und Arg87 größere, Tyr130 und Lys173 immerhin deutliche strukturelle Unterschiede aufweisen. Tyr176, Tyr420, Arg172 und Arg162 dagegen, aber auch Lys112 und Lys124 sind strukturell konstant. Kann für die ersteren, flexiblen Bereiche der Domäne I angenommen werden, dass sie sich möglicherweise im Verlaufe der DNA-Bindung in Richtung des Bindungspartners bewegen, so könnten letztere, strukturell unflexible Reste an der initialen Positionierung der DNA beteiligt sein. Für diese Subdomäne wird aufgrund eines geometriebasierten Docking-Modells eine Flip-Bewegung für möglich gehalten, die etwa 60° relativ zum Kernbereich der Domäne I beträgt. Vier definitive Kontaktstellen jeweils von PI-SceI und der Erkennungssequenz wurden strukturell überlagert. Es handelt sich um Asp218 und C⁻³(unterer Strang), Asp326 und G⁺³ (oberer Strang), Tyr328 und G⁺⁴ (unterer Strang) sowie His333 und T⁺⁹ (oberer Strang), also um die endonukleolytischen Zentren und zwei Photo-Quervernetzungen ohne Abstandshalter. Das resultierende Modell führt zu einer Orientierung der DNA relativ zur Domäne I, die den biochemischen Daten und der Verteilung der Oberflächenladung widerspricht. Um in Übereinstimmung des Modells mit den biochemischen Daten zu kommen, wird eine Rotationsbewegung der zangenartigen Subdomäne (oder Teile davon) vorgeschlagen, die die Stränge **i**, **j** und **k** in Kontakt mit der großen Furche im Bereich der Basenpaare +16 bis +18 bringt sowie die Aminosäuren Arg90/Arg94 mit A¹⁴/T¹⁵/T¹⁶ des oberen Stranges und Lys170/Lys173 mit Nukleotiden 10 bis 12.

6.2. Englisch - English

6.2.1. Nicotinamide mononucleotide adenylyltransferase of *Homo sapiens* in complex with nicotinamide mononucleotide

Nicotinamide adenin dinucleotide (NAD⁺) is an important molecule as coenzyme in cellular redox reactions and signal-transduction pathways. The enzyme nicotinamide/nicotinate mononucleotide adenylyltransferase (NMNAT/NaMNAT) takes part in the biosynthesis of NAD⁺ by transferring the adenosine monophosphate moiety of ATP to nicotinamide/nicotinate mononucleotide. In the case of NaMN the product is nicotinate-adenin dinucleotide (NaAD⁺) which gets amidated by NAD⁺-synthetase. In the case of NMN, NAD⁺ gets synthesized directly.

The crystal structure of NMNAT from *Homo sapiens* in complex with NMN was solved to a maximum resolution of 2.9 Å with the SAD method (single wavelength anomalous dispersion). Final R_{cryst} and R_{free} values are 24.6 % and 28.6 % respectively. The structure of NMNAT consists of a six-stranded parallel β-sheet with helices on both sides, which in the core is the Rossmann fold. Electron density was observed for the ligand NMN but not for a loop of 37 amino acids, 109 to 146, that is positionally disordered. The structure of hNMNAT differs from the homologous proteins of *M. thermoautotrophicum* and *M. jannaschii* by this loop that contains a nuclear localization signal and by additional amino acids that form the helices **H** and **I** and the strand **f** in the human enzyme. Those secondary structure elements cause a different oligomerization compared to the archaeal proteins, but all occur as hexamer as biological unit. Protein-protein contacts between the monomers that form two trimers on top of each other are established between helices **A**, **H** and **J** as well as the loop between Helices **J** and **K**. The trimers are twisted with respect to each other, the β-sheets of subunits of different trimers arrange side by side in an antiparallel fashion and form a kind of a super-secondary structure but do not contact directly. Bacterial NMNAT occurs as monomer (*E. coli*) or dimer (*B. subtilis*) and in the case of *E. coli* NMNAT has further secondary structure elements.

The ligand NMN is bound by the amino acids Ser16, Lys57, Trp92, Thr95, Leu168 and Trp169, all on one side of the β -sheet. Comparison of the NMN complex with human NMNAT structures solved by other groups (the apoenzyme and the NAD⁺ complex) and with the archaeal and procarial homologs yield a model for ligand binding and synthesis. In the apoform, the enzyme is in an open state. Helices **F** and **G** on one side and the region around helix **B** on the other side have a greater distance than in the complex. NMN binds to this part of the ligand binding pocket, on one side of the β -sheet. ATP, the other reaction partner, is possibly attracted by a cluster of positively charged amino acids that only forms upon oligomerization. It could get transferred to the reaction center, possibly under participation of amino acids Lys58 and Arg231. ATP binds on the other side of the β -sheet. Binding is facilitated by Phe17, Gly156, Glu215 and Asn219 and most likely by the conserved residues Thr21, His24, Ser222, Thr224 and Arg227 which obviously bind the β - and γ -phosphates. Upon substrate binding, the above mentioned elements approach and at the same time the C-terminus gets ordered (from residue 258 on). It covers parts of the NMN-binding pocket and of His24 and Lys58. Amino acids Gly15, Ser16, Phe17, His24 and Lys57 might participate in the synthesis reaction because they are locally close to the phosphodiester bond that is going to be formed. Dissociation of NAD⁺ and pyrophosphate PP_i goes along with relaxation of the protein and completes the reaction cycle.

6.2.2. Domain I of the homing endonuclease PI-SceI of *Saccharomyces cerevisiae*

The homing endonuclease PI-SceI of *S. cerevisiae* is an intein, an *internal protein*, embedded into the extein sequences of the 69 kDa subunit of the vacuolar membrane H⁺-ATPase. In a protein splicing process it cuts itself out of the protein precursor and combines the two exteins. Responsible is domain I of PI-SceI that at the same time accounts for most of the binding energy to the specific DNA sequence of at least 31 bp. The recognition sequence is the *VMA1-Dvde* locus of *S. cerevisiae* chromosome 8. This locus is the allele of the *VMA1* genetic sequence (VMA: vacuolar membrane H⁺-ATPase) that is deficient for *vde* (VDE: VMA derived endonuclease = PI-SceI). Domain II of PI-SceI contains the nucleolytic center, which cuts the specific recognition sequence and initiates the insertion of the *vde*-gene ("homing"). It is a member of the LAGLIDADG family of homing endonucleases and binds the DNA only loosely.

The crystal structure of domain I of PI-SceI was solved with the molecular replacement method to a maximal resolution of 1.35 Å. Final R_{cryst} and R_{free} values are 15.0 % and 18.9 % respectively. Although four crystal structures of the whole protein were known already, the high resolution and the good definition of areas that were not or not well defined in the other structures give a view of further details. The goals to observe the whole protein sequence in the electron density and to produce protein-DNA cocrystals were not reached. Rather, protein crystals were observed in crystallization attempts with complexes. 96 of 217 amino acids take part in crystal contacts. Domain I of PI-SceI, or the core part of it, adopts the hint fold of the hedgehog/intein domain and consists mainly of β-strands. The active center of domain I, the protein splicing site, has cysteine 1 and two N-terminal extein residues but lacks the C-terminus of the intein, Asn454. The common model for the first step of the protein splicing reaction mechanism can be structurally confirmed. The side chain of Cys1 initiates the initial N-S acyl shift by a nucleophilic attack of the carbonyl carbon of Ala(-1). Subsequently follow a transesterification, formation of a succinimide and the final S-N acyl shift.

Residues Gln55 to Glu66 were not observed in the electron density due to local disorder. They obviously form a flexible loop that is suspected to take part in the tight binding of DNA by domain I. Further residues that are supposed to

participate in DNA binding because of biochemical data, lie at the front side of a tongs-like subdomain that is formed by strands *i*, *j* and *k* and helices **B** and **C**. Most likely, helices **D** and **E** as well as strand *l* are also part of this subdomain because they also contain potential DNA-binding residues. A structural alignment of all known protein chains of PI-SceI domain I suggests, that besides the flexible residues Gln55 to Glu66 the amino acids Arg94, Arg90 and Arg87 have bigger, Tyr130 and Lys173 at least significant structural differences. Tyr176, Tyr420, Arg172 and Arg162, but also Lys112 and Lys124 on the other side are structurally constant. For the first, flexible group it can be suggested, that they possibly rearrange to facilitate DNA binding, for the latter, inflexible residues that they might take part in the initial positioning of the DNA.

A geometry based docking model makes the possibility of a movement of about 60° relative to the core visible. Four proven contact sites of PI-SceI and the recognition sequence were structurally aligned. Those are Asp218 and C⁻³(bottom strand), Asp326 and G⁺³ (top strand), Tyr328 and G⁺⁴ (bottom strand) as well as His333 und T⁺⁹ (top strand), that means the nucleolytic active sites and two zero-length photo-crosslinks. The resulting model yields an orientation of the DNA relative to domain I that is not consistent with biochemical data and the distribution of surface charges. To reach agreement of biochemical data and the model, a rotational movement of the tongs-like subdomain (or parts thereof) was suggested that brings the strands *i*, *j* and *k* to close contact with the major groove of basepairs +16 to +18, the amino acids Arg90/Arg94 close to A¹⁴/T¹⁵/T¹⁶ of the top strand and Lys170/Lys173 close to the nucleotides 10 to 12.