Strukturelle Untersuchungen an Varianten des *Ecballium elaterium* Trypsin Inhibitors-II (EETI-II)

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Ralph Krätzner
aus Fulda

Göttingen 2001

D 7

Referent: Prof. G. M. Sheldrick, Ph. D.

Korreferent: Prof. Dr. P. Susse

Tag der mündlichen Prüfung: 27.6.2001

Inhaltsverzeichnis

1 EINLEITUNG 1

1.1 Das Inhibitor-Cystinknoten-Motiv (ICK) 2

1.2 Knoten in Proteinen 3

1.3 Verbreitung des Inhibitor-Cystinknoten-Motivs 4

1.4 Strukturuntersuchungen an Cystinknoten-Proteinen 9

2 AUFGABENSTELLUNG 13

3 MATERIALIEN UND METHODEN 14

3.1 Biologische Materialien 14

3.2 Molekularbiologische Methoden 16

3.2.1 Stammhaltung und Vermehrung von *E. coli* Stämmen 16

3.2.2 Vorbereitung von Geräten und Lösungen 17

3.2.3 Transformation von *E. coli* nach der Elektroporationsmethode 17

3.2.4 Spaltung von Proteinen an Methionresten mit Bromzyan 18

3.2.5 Denaturierende Polyacrylamidgelelektrophorese 19

3.2.6 Reversed Phase High Performance/Pressure Liquid Chromatography 20

3.2.7 Affinitätschromatographie über immobilisierte Metallionen *(IMAC)* 21

3.2.8 Bestimmung der Ausbeute an EETI-II 21

3.3 Kristallographische Methoden 22

3.3.1 Umgang mit Proteinlösungen zu kristallographischen Zwecken 22

3.3.2 Kristallisation von Proteinen 22

3.3.3 Kristallmontage 24

3.3.4 Röntgenquellen 26

3.3.5 Flächenzähler 27

3.3.6 Datensammlung 29

3.3.7 Datenprozessierung 30

3.3.8 Strukturlösung 31

3.3.9 Strukturverfeinerung 37

3.3.10 Modelldarstellung 38

3.3.11 Generierung von 3D-Bildern 39

3.3.12 Evaluierung von Strukturen 39

4 ERGEBNISSE 40

4.1 Aufreinigung von EETI-Varianten zu kristallographischen Zwecken 40

4.2 Vorarbeiten zu Röntgenstrukturbestimmung von EETI-II (wt) 44

4.3 Ansätze zu Strukturlösung von EETI-II (wt) 47

4.4 Modelle für die Packung von EETI-II (wt) 52

4.5 Versuch der Strukturlösung von EETI-II (wt) über anomale Dispersion 53

4.6 Strukturbestimmung des Trypsin/EETI-II (wt) Komplexes 55

4.7 Strukturbeschreibung des Trypsin/EETI-II (wt) Komplexes 61

4.8 Varianten von EETI-II im Komplex mit Trypsin 67

4.9 Strukturbestimmung des Trypsin/EETI^{TNNK} Komplexes 68

4.10 Strukturbeschreibung des Trypsin/EETI^{TNNK} Komplexes 72

4.11 Strukturbestimmung und Beschreibung der Komplexe EETI^{NEDE} und EETI^{ERDE} 74

4.12 Vergleich der Ca-Atomkoordinaten von Trypsin/EETI-II (wt) und Trypsin/EETI^{TNNK} 76

4.12.1 Trypsin 76

4.12.2 Vergleich der Strukturen von EETI-II (wt) und EETI^{TNNK} 77

4.13 Vergleich von EETI-II (wt) mit den Strukturen von CMTI-I und MCTI-A 80

4.14 Vergleich der EETI-II (wt) Röntgenstrukturen und der NMR Struktur .. 82

5 DISKUSSION 83

5.1 Zur Funktion des #22-#24 b-turns 86

6 ZUSAMMENFASSUNG 89

7 LITERATUR 91

Abkürzungsverzeichnis

Abb.	Abbildung
Ala	Alanin
APS	Ammoniumperoxodiphosphat
Arg	Arginin
Asn	Asparagin
Asp	Aspartat
Bicine	N,N-bis(2-Hydroxyethyl)-glycin
CCD	*charged coupled device*
c-terminal	carboxy-terminal
DESY	Deutsches Elektronen Synchrotron
DNA	*deoxyribonucleic acid*
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Dinatriumethylendiamintetraacetat
EETI-II	*Ecballium elaterium* Trypsin Inhibitor-II
EMBL	*European Molecular Biology Laboratory*
ESRF	*European Synchrotron Radiation Facility*
et al.	*et alii* (lat.:und andere)
h	Stunde
HPLC	*high performance/pressure liquid chromatography*
kDa	Kilodalton
M	mol/l
min	Minute
MWPC	*multiwire proportional chamber*

n-terminal	amino-terminal
OD_{600}	Optische Dichte bei der Wellenlänge 600 nm
PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerase-Ketten-Reaktion
Pro	Prolin
PVC	Polyvinylchlorid
RP	reversed phase (Umkehrphasen)
r.m.s.	*root mean square* (mittlere quadratische Abweichung)
UpM	Umdrehungen pro Minute
RT	Raumtemperatur
SDS	Natriumdodecylsulfat
Ser	Serin
Tab.	Tabelle
TEA	Tri-Ethanolamin
TEMED	N,N,N‘,N‘-Tetramethylendiamin
TFA	Trifluoressigsäure
Tris	Tris-(hydroxymethyl)-aminoethan
UV	ultraviolett
Ü. N.	über Nacht
Vol.	Volumen
(v/v)	Volumen zu Volumen
(wt)	Wildtyp
(w/v)	Masse zu Volumen
(w/w)	Masse zu Masse
x g	-fache Erdbeschleunigung

1 Einleitung

Proteine sind die unter allen Aspekten der Chemie vielseitigsten bekannten Polymere. Auf der Basis von 20 unterschiedlichen „Bauelementen“, den natürlichen Aminosäuren, wurde im Laufe der Evolution eine bis heute unüberschaubare Zahl an Makromolekülen mit diversen Funktionen und Eigenschaften hervorgebracht. Proteine übernehmen unter anderem strukturelle, immunologische, enzymatische, rezeptorische, regulatorische und inhibitorische Aufgaben:

Beispielsweise wird die Transparenz der Augenlinse von Wirbeltieren durch die besonderen Eigenschaften der Crystalline (Norledge *et al.* 1997) hervorgerufen. Diese Proteine müssen über einen langen Zeitraum stabil bleiben, da der Stoffwechsel innerhalb der Linse sehr eingeschränkt ist. Collagene sind für die Stabilität von Knochen, Sehnen und Bändern verantwortlich. Antikörper identifizieren im Verlauf der Immunantwort des Körpers hochspezifisch Antigene und ermöglichen so eine effiziente Abwehr von Krankheitserregern.

Sehr weit diversifiziert sind die enzymatischen Funktionen von Proteinen: ATPasen von Pflanzen und Bakterien erzeugen über Membranpotentiale chemische Energieträger wie ATP, das von anderen Enzymen als Energielieferant für den Zellstoffwechsel umgesetzt wird. Eine Vielzahl von Proteinen erfüllt regulatorische Aufgaben. Der lac-Repressor steuert über Bindung an die DNS die Expression des β-Galaktosidase-Gens. Bei der Lichtrezeption auf der Netzhaut von Säugetieraugen ist Rhodopsin beteiligt. Bei vielen Enzymen sind nur relativ wenige Aminosäurereste direkt an der katalytischen Umsetzung eines Substrates beteiligt. Im Falle der 223 Aminosäurereste umfassenden Serinprotease Trypsin wird das katalytische Zentrum von nur drei Aminosäuren gebildet (Otlewski & Zbyryt 1994), der Rest der Sequenz stellt das funktionstragende, strukturelle Rahmenwerk dar.

Demgegenüber weisen einige Proteine, die hormonelle oder inhibitorische Funktionen wahrnehmen, eine im Durchschnitt relativ geringe Größe auf. Das aus der Bauchspeicheldrüse isolierte Hormon Insulin umfasst zwei Ketten mit insgesamt nur 51 Aminosäuren (Hodgkin 1974, Gursky *et al.* 1993).

In dieser Größenordnung stellt sich die Frage, wo die Trennlinie zwischen Proteinen und Polypeptiden gezogen werden soll. Ein wichtiges Kriterium hierfür kann die Ausbildung einer definierten Tertiärstruktur sein: Proteine müssen eine Faltung

aufweisen, während Polypeptide keine Vorzugskonformation im Hauptketten-Verlauf einnehmen.

1.1 Das Inhibitor-Cystinknoten-Motiv (ICK)

Eine zentrales Strukturelement, das bei einer Reihe relativ kleiner Proteine die Faltung im konformationellen Sinne nach oben genanntem Strukturkriterium ermöglicht, ist das sogenannte Inhibitor-Cystinknoten-Motiv (ICK), das durch 3 Cystin-Brücken gebildet wird: Die Disulfidbrücken, deren Verknüpfung mit dem Polypeptidrückgrat an ein Knotengebilde erinnert, bilden das strukturelle Rahmenwerk für eine im Eukaryontenreich weit verbreitete Gruppe kleiner Proteine, die fast ausschließlich inhibitorische Funktionen erfüllen (Norton *et al.* 1998). Die hydrophobe Kernregion größerer Proteine wird bei diesen nur ca. 30-50 Aminosäuren umfassenden Naturstoffen durch die Fixierung der Polypeptidkette über sechs typisch verknüpfte und in ihrer Sequenzposition hoch konservierte Cysteinreste ersetzt. Abb. 1 zeigt ein Modell des Conotoxins MVIIA als Beispiel für ein Cystinknoten-Protein . Das erste Cystein der Peptidsequenz ist dabei mit dem 4., das 2. mit dem 5. und das 3. mit dem 6. Cystein verknüpft. Zwei Cystinbrücken bilden dabei mit der Hauptkette einen Ring, der von der dritten Disulfidbrücke durchdrungen wird. Die Inhibitor-Cystinknoten-Familie weist als zusätzliches Charakteristikum ein kurzes, dreisträngiges, antiparalleles β-Faltblatt und im Falle der *squash*-Inhibitoren eine kurze $^{3}{}_{10}$ Helix auf. Isolierte Cystinknoten treten auch in größeren Proteinen auf, so z. B. im 114 Aminosäurereste umfassenden *Human Bone Morphogenetic Protein-2* (BMP-2) (Scheufler *et al.* 1999).

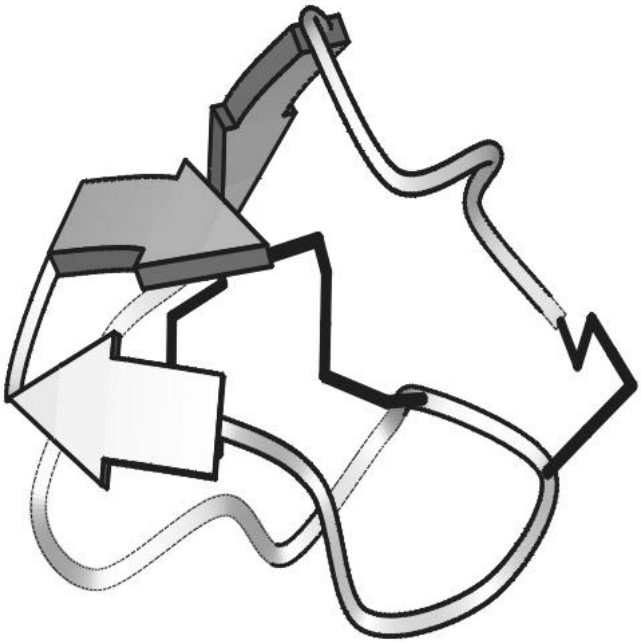


Abb.1: Darstellung des w-Conotoxins MVIIA als Beispiel für ein Cystinknoten-Protein. Die Cystinbrücken sind schwarz eingezeichnet (Koordinaten nach Kohno *et al.* 1995)

1.2 Knoten in Proteinen

Beim Cystinknoten-Motiv handelt es sich nicht um einen Knoten im mathematischen Sinne. Abb. 2 zeigt am Beispiel des Trypsininhibitors EETI-II, dass der „Cystinknoten“ mit Hilfe der in der Knotentheorie entwickelten Reidemeister-Züge (Beliakova & Schumakovitch 2000) auf ein einfaches Gebilde reduziert werden kann. Die gezielten Untersuchungen des Peptid-Rückgrats von 3440 Proteinen aus der PDB ergaben, dass insgesamt nur acht Strukturen existieren, die einen echten Knoten aufweisen (Taylor 2000). Dabei fallen diese Knoten meist unter die einfachste Klassifikation im mathematischen Sinne, das sogenannte Trefoil. Eine einzige Struktur, die der Hydroxy-Essigsäure-Isomerase (Biou *et al.* 1997, PDB Identifikation 1yve und 1qmg) weist einen komplizierteren Knoten (*figure of eight*) auf. Die Knotenbildung erfolgt hier sehr wahrscheinlich durch posttranslationale Modifikation.

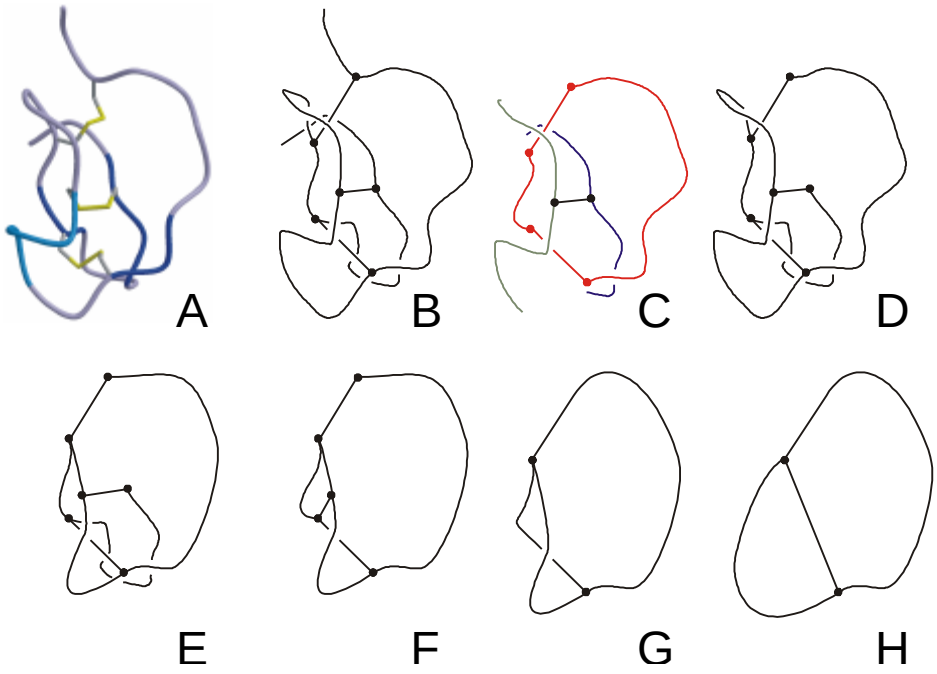


Abb.2: Die „Entknotung“ des Cystinknotens kann mit wenigen erlaubten Operationen durchgeführt werden. Die Struktur eines ICK-Proteins (A) lässt sich als Strichdiagramm (B) darstellen. Aus (C) wird deutlich, dass das Peptidrückgrat des Proteins mit zwei Disulfidbrücken einen Ring (rot) bildet, der von der dritten Cystinbrücke (schwarz), die einen darüber (grün) und einen darunter verlaufenden (blau) Strang verbindet, durchdrungen wird (Diagramm (C) wurde vereinfacht dargestellt). In einem ersten Schritt werden die terminalen Enden der Polypeptidkette bis auf die nächsten T-Verzweigungen gekürzt (D), die Schleifen zwischen den Verzweigungen werden zu Zwecken der Übersichtlichkeit verkürzt (E und F). Parallele Verzweigungen dürfen nach der Knotentheorie wegfallen (F). Das resultierende Gebilde (G) weist übersichtlich dargestellt (H) keine Ähnlichkeit zu einem Knoten auf.

1.3 Verbreitung des Inhibitor-Cystinknoten-Motivs

Auf der Basis des stabilen strukturellen Rahmenwerkes des Inhibitor-Cystinknoten-Motivs wurde eine große Anzahl konformationell eingeschränkter Peptide evolviert, die bei relativ geringer Molmasse Träger unterschiedlichster biologischer Spezifitäten sind. Eine wichtige Funktionalisierung des Strukturgerüstes tritt beispielsweise bei den Conotoxinen auf: Schnecken der Gattung *Conus* sind als marine Jäger über die ganze Welt verbreitet und haben eine hoch spezialisierte Strategie entwickelt, ihre Beute zu vergiften und bewegungsunfähig zu machen. Einige fischjagende Arten, deren Habitate vor allem im Indo-Pazifischen Raum lokalisiert sind, locken ihre Beute durch einen beweglichen Stechrüssel an, der einen Wurm vortäuscht. Im Moment des Zustechens

bei ausreichender Nähe zum Opfer wird der Inhalt einer Giftdrüse unterhalb des Rüssels injiziert, wodurch der Fisch sofort paralysiert wird (Terlau *et al*. 1996).

Das Venom dieser Schnecken enthält je nach Art 40-200 verschiedene kleine, konformationell eingeschränkte Peptide diverser Funktionen, die allesamt eine sehr spezifische biologische Aktivität aufweisen. Die meisten der 10-30 Aminosäuren großen Toxine lassen sich ihrer Wirkungsweise nach in vier Hauptgruppen einteilen: ω-Conotoxine, die ca. 30 Aminosäuren umfassen, zeigen das Strukturmotiv der Inhibitor-Cystinknoten-Familie. Während δ-Conotoxine bei Rezeptorbindung die Durchlässigkeit spannungsregulierter Natrium-Kanäle erhöhen, binden ω-Conotoxine mit hoher Selektivität an Calcium-Kanäle, deren Durchlässigkeit für die Ionen dadurch drastisch gesenkt wird (Olivera *et al.* 1984, Shon *et al.* 1994, Shon *et al.* 1995). μ-Conotoxine, die ebenfalls drei Disulfidbrücken ausbilden, unterbinden den Ionenfluss durch bestimmte Natriumkanäle, wobei sie allerdings nicht als direkte Antagonisten zu den δ-Conotoxinen auftreten, da sie an andere Subtypen binden (Olivera *et al.* 1985, Olivera *et al.* 1995). Die Gruppe der α-Conotoxine umfasst kleine Peptide von ca. zehn Aminosäuren Länge, die durch zwei Disulfidbrücken fixiert sind und als Inhibitoren von Acetylcholin-Rezeptoren wirken (Shon *et al.* 1994). Andere, seltenere Vertreter dieser Naturstoffe werden als κ-Conotoxine zusammengefasst. Bei allen ω, δ, μ und κ-Conotoxinen tritt die gleiche Konnektivität der Cysteinreste auf. Versuche mit Mäusen, denen einzelne Fraktionen aus dem Toxingemisch der Kegelschnecken injiziert wurden, lieferten deutliche Hinweise darauf, dass einige Peptide Träger von pharmazeutisch hochinteressanten Funktionalitäten sind (Olivera *et al.* 1995). Als erstes Schneckengift steht Ziconotide (SNX-111), ein synthetisches Produkt, das sich vom ω-Conotoxin MVIIA aus dem Venom der Schnecke *Conus magnus* ableitet, unmittelbar vor der Markteinführung durch die kalifornische Firma Neurex Corp. Wichtigstes Anwendungsgebiet dieses Medikamentes ist die Schmerztherapie im Falle von fortgeschrittenen AIDS- oder Krebserkrankungen. Darüber hinaus soll Ziconotide zur Behandlung von Neurodegeneration als Folge einer Ischämie bei Kopf-Traumata eingesetzt werden.

Eine weitere Gruppe von Tieren, aus der Cystinknoten-Proteine isoliert werden konnten, umfasst die australischen Trichternetz-Spinnen (engl. *funnel web spider*). Das Neurotoxin Robustoxin der Spinne *Atrax robustus* beispielsweise besitzt 42 Aminosäuren, drei der vier Cystinbrücken bilden das Knotenmotiv (Temple *et al*.

1999). Die Spinnentoxine können ebenso wie die Conotoxine letale Folgen beim Menschen haben und dienen zum Beutefang und zur Verteidigung. Abb.3 zeigt die Kegelschnecke *Conus textilis*.

Abb.3: Bild der Kegelschnecke *Conus textilis*. Der Stechrüssel der ca. 10 cm langen Schnecke tritt unterhalb eines zum Atmungsorgan gehörenden, rohrartigen Fortsatzes aus dem Gehäuse des Mollusken aus. Auf Beutefang lockt die Schnecke, im Meersand vergraben, mit diesem Rüssel Fische an, die bei ausreichender Nähe harpuniert werden. Photo: Dr. Bruce G. Livett & Mr. David Paul[1].

Inhibitor-Cystinknoten-Proteine konnten auch aus einer Vielzahl von Pflanzen isoliert werden. Das zirkuläre Protein Kalata B1 wird in der Pflanze *Oldenlandia affinis* produziert (Daly & Craik 2000), die bei Naturvölkern aufgrund ihrer Uterus-kontrahierenden Wirkung zur Geburtshilfe eingesetzt wird. Gurmarin ist ein Cystinknoten-Protein, das aus der indischen Pflanze *Gymnema sylvestris* stammt und den süßen Geschmackssinn inhibiert. Es weist im Gegensatz zu dem oben beschriebenen, dreisträngigen nur ein zweisträngiges antiparalleles β-Faltblatt auf. Andere Substanzen aus den Blättern dieser Pflanze unterdrücken den sauren Geschmack, wodurch der Verzehr von *Gymnema sylvestris* eine Verwirrung der Schmeckrezeption hervorruft. Aufgrund dieses Effektes, der beim fressenden Tier zu Appetitverlust führt, werden Blätter von *Gymnema sylvestris* als Tee zur Diäthilfe gehandelt (Fletcher *et al.* 1999, Ninomiya & Imoto 1995).

Eine andere wichtige Gruppe von Cystinknoten-Proteinen umfasst die sog. *squash*-Inhibitoren von Trypsin. Die *squash*-Inhibitoren wurden aus den Samen von Kürbisgewächsen isoliert und weisen die selben Strukturcharakteristika wie die bereits besprochenen Cystinknoten-Proteine auf, allerdings bilden sie zusätzlich noch eine kurze $^{3}{}_{10}$ Helix als Sekundärstrukturmotiv aus. Sie gehören mit Dissoziationskonstanten im nanomolaren Bereich zu den stärksten bekannten Inhibitoren von Serin-Proteasen wie Trypsin und sind mit einer Länge von durchschnittlich ca. 30 Aminosäuren etwa

[1]Dr. Bruce G. Livett, Dept. Biochemistry and Molecular Biology, University of Melbourne, Australia, and Mr. David Paul, Dept. Zoology, University of Melbourne, Australia.

halb so groß wie Trypsininhibitoren vom Kasal- oder Kunitz-Typ (BPTI bzw. PSTI, Übersicht bei Laskowski & Kato 1980). Die Interaktion mit Trypsin erfolgt hauptsächlich durch die Inhibitorschleife, die die sechs Aminosäurereste zwischen den ersten beiden Cysteinen umfasst und die bei allen bekannten *squash*-Inhibitoren weitgehend konserviert sind. Bei Serinprotease-Inhibitoren ist der erste Aminosäurerest dieses inhibierenden *loops* immer ein Prolin, der zweite meistens ein Arginin, allerdings kann diese sogenannte P1-Position auch durch ein Lysin besetzt sein. Der dritte Rest ist immer ein Isoleucin, in den Positionen darauf folgen häufig Leucin (#4), Methionin (#5) und Glutamat (#6). Die P und S Benennungen von Aminosäure-Positionen in der Inhibitorschleife geht auf die von Schechter & Berger (1967) eingeführte Nomenklatur zurück und ist bis heute gebräuchlich. In dem 1967 vorgeschlagenen Modell binden die Aminosäurereste des Polypeptid-Substrats an sogenannte *subsites* im als *active site* bezeichneten aktiven Zentrum der Protease. Diese *subsites* werden entsprechend mit S bezeichnet, die Aminosäurereste des Substrats mit P (*peptide*). Abb.4 verdeutlicht dieses Schema.

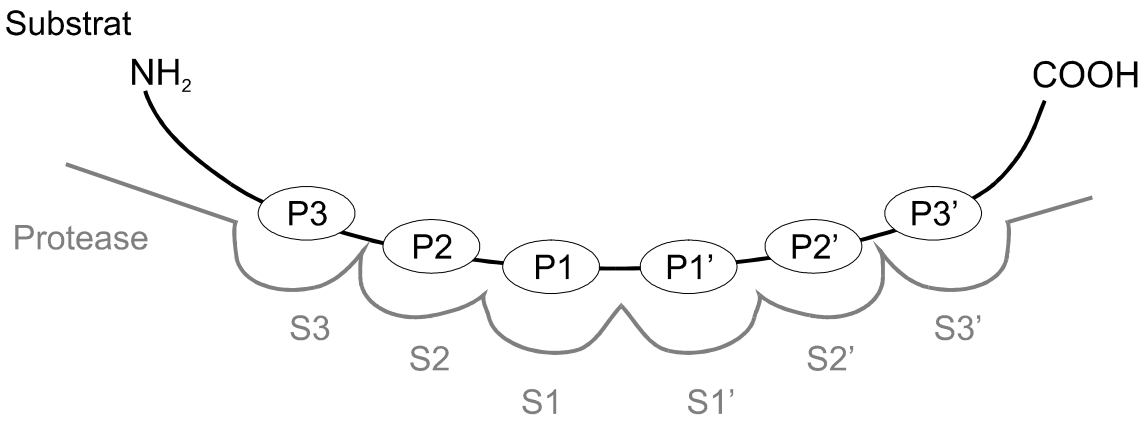


Abb.4: Schematische Darstellung der Substrat/Protease-Interaktion nach Schechter & Berger 1967. Die Aminosäurereste des Substrats binden an sogenannte *subsites* des aktiven Zentrums der Protease. Die Reste auf der n-terminalen Seite der Schnittstelle werden mit P1, P2,P3, die auf der c-terminalen Seite mit P1‘, P2‘, P3‘ bis P8‘ bezeichnet. Die Bezeichnung für die *subsites* erfolgt analog. Die Bindung, an der das Substrat gespalten wird, ist nach dieser Nomenklatur die zwischen P1 und P1‘.

Die Spaltung des Substrats erfolgt an der Bindung zwischen P1 und P1‘. Auf der n-terminalen Seite der Protease-Schnittstelle werden die Aminosäurereste des Peptids mit P1, P2 und P3, die auf der c-terminalen Seite mit P1‘, P2‘, P3‘ bis zu P8‘ bezeichnet.
Die entsprechenden *subsite*-Positionen der Protease werden analog benannt.

Durch die Bindung des Inhibitors an Trypsin, bei der in Komplexen dieses Typs mehr als 50% der Wechselwirkungen von der P1-Position des Inhibitors ausgehen (Bode *et al.* 1989), wird die katalytische Triade blockiert, wodurch die proteolytische Aktivität gehemmt wird (Otlewski & Zbyryt 1994). In Abb. 5 sind einige exemplarische Peptidsequenzen für ICK-Proteine aufgeführt:

```
CMTI I        RVCPRILMECKKDSD-CLAECV---CLEHGY-----C
Gurmarin      QQCVKKDELCIPYYLDC---CEPLECKKVNWWDHK-CIG
EETI II        GCPRILMRCKQDSD-CLAGCV---CGPNGF-----CG
GVIA             CKSOGSSCSOTSYNC---CRS—-CNOYTKR----CY
MVIIA            CKGKGAKCSRLMYDC---CTGS-CRSGK------C
Robustoxin       CAKKRNWCGKNED-CC--CPMK-CIYAWYNQQQSCQTTITGLFKKC
SVIA             CRSSGSOCGVTSI-C---CGR--CYRGK------CT
TXVIA           WCKQSGEMCNLLDQNC---CDGY-CIVLV------CT
GMVIA        VKPCRKEGQLCDPIFQNC---CRGWNCVLF-------CV
Kalata B1       TCVGGT--CNTPG—-CT--CSWPVCTRNGLPV---CGE
```

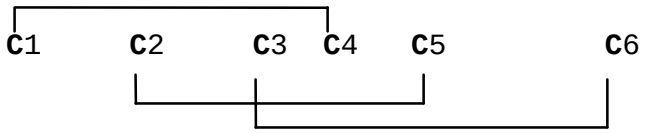


Abb. 5: Sequenzvergleich von Mikroproteinen, die das "Inhibitor-Cystinknoten-Motiv" aufweisen. Die Cysteine sind fett gedruckt. Unterhalb befindet sich ein Schema für die Verknüpfung der Cystinbrücken. CMTI-I: Trypsininhibitor aus *Curcubita maxima*; Gurmarin: Geschmacks-Supressor aus *Gymnema sylvestris*; EETI-II: Trypsininhibitor aus *Ecballium elaterium*; GVIA: w**-Conotoxin aus *Conus geographus*; MVIIA:** w**-Conotoxin aus *Conus magnus*; Robustoxin: Neurotoxin aus *Atrax robustus*; SVIA:** w**-Conotoxin aus *Conus striatus*; TXVIA:** d**-Conotoxin aus *Conus textilis*; GMVIA:** d**-Conotoxin aus *Conus glomaris*; Kalata B1: zirkuläres Protein aus *Oldenlandia affinis* (modifiziert nach Olivera 1995, Otlewski 1990).**

Die Natur hat auf der Basis des ICK-Motivs offenbar eine Reihe hochwirksamer Inhibitoren unterschiedlichster Funktion evolviert, von denen einige hohe pharmakologische Relevanz haben könnten. Ein generelles Problem in der Therapie durch Medikamente auf Proteinbasis liegt in ihrer komplizierten klinischen Anwendbarkeit. So muss beispielsweise das beschriebene MVIIA-Derivat mittels einer speziellen, implantierten Pumpe direkt an seinen Wirkungsort im menschlichen Organismus befördert werden. Insulin als weiteres Beispiel muss bislang mittels Injektion appliziert werden. Darüber hinaus ist die industrielle Darstellung von Polypeptid-Drogen häufig ein ökonomisches Problem, insbesondere führt die Synthese der Conotoxine aufgrund seltener Aminosäuren in der Polypeptidkette und der komplizierten *in vitro* Faltung oftmals zu geringen Ausbeuten (Price-Carter *et al.* 1996). Eine Strategie, diese Schwierigkeiten zu umgehen, ist das strukturbasierte Design neuer

Medikamente. Im Prinzip sollen hierbei die funktionellen Eigenschaften der jeweiligen Proteine oder anderer problembehafteter Liganden auf einfache und im großen Maßstab produzierbare organische Moleküle übertragen werden (Klebe 2000). Eine Schlüsselrolle kommt hierbei der exakten Bestimmung der räumlichen Struktur des jeweiligen Makromoleküls zu, die als sogenannte Leitstruktur für die Übertragung der Wirksamkeit auf eine zu synthetisierende kleinere Verbindung herangezogen werden kann.

Zur Bestimmung einer solchen Leitstruktur auf Proteinbasis sind zwei strukturchemische Untersuchungsmethoden wichtig: 2D-NMR-Experimente liefern allgemein ein sehr gutes Bild der Moleküldynamik in Lösung. Die Röntgenstrukturanalyse aus hochaufgelösten Datensätzen lässt dagegen wesentlich genauere Aussagen über die dreidimensionale Geometrie, Torsionswinkel und intra- sowie intermolekulare Wechselwirkungen zu. Beide Analysemethoden ergänzen sich in wesentlichen Punkten und können gemeinsam die notwendige Information für die Übertragung der Funktionalität eines der beschriebenen Naturstoffe auf ein anderes Molekül zur Verfügung stellen.

1.4 Strukturuntersuchungen an Cystinknoten-Proteinen

Strukturelle Untersuchungen an ICK-Proteinen sind bisher vor allem mit NMR-Methoden durchgeführt worden. Im Bereich der Conotoxine sind eine Reihe von Strukturen in der Protein Datenbank (PDB) **(**Berman *et al.* 2000) hinterlegt, beispielsweise für das ω-Conotoxin MVIIA (Kohno *et al.* 1995), das κ-Conotoxin PVIIA (Savarin *et al.* 1998), das μ-Conotoxin GIIIA (Wakamatsu *et al.* 1992) und für das δ-Conotoxin TXVIA (Kohno *et al.* to be published, PDB-ID 1FU3). Auch die Strukturen der erwähnten Cystinknoten-Proteine Kalata B1 (Craik & Norman, 1995) und Gurmarin (Fletcher *et al.* 1999) wurden mittels NMR-Methoden aufgeklärt, dasselbe gilt für die Gifte der Trichterspinnen. Aus der Familie der *squash*-Inhibitoren wurden einige Vertreter im Komplex mit Trypsin kristallisiert, daneben wurden zahlreiche NMR- Strukturen bestimmt. CMTI-I (*Cucurbita maxima* Trypsin Inhibitor-I) wurde im Komplex mit Trypsin aus verschiedenen Organismen kristallisiert und röntgenkristallographisch untersucht (Bode *et al.* 1989; Helland *et al.* 1999), daneben wurde die NMR-Struktur von CMTI-I in Lösung bestimmt (Nigles *et al.* 1991). Von

Cucurbita pepo Trypsin Inhibitor (CPTI) ist ein Komplex mit Trypsin aus dem Rinderpankreas beschrieben (Helland *et al.* 1999). Röntgenstrukturanalysen wurden darüber hinaus an Komplexen von MCTI-A (*Mamordica charantia* Trypsin Inhibitor-A) mit β-Trypsin aus dem Schweine- (Huang *et al.* 1993) und aus dem Rinderpankreas (Zhu *et al.* 1999) durchgeführt. Der vielleicht am besten untersuchte Vertreter der Familie der *squash*-Inhibitoren ist der *Ecballium elaterium* Trypsin Inhibitor-II (EETI-II) aus der „Spritzgurke" *Ecballium elaterium* (Abb.6). Von EETI-II und Varianten wurden Strukturen mittels 2D-NMR ermittelt (Heitz *et al.* 1989), die Röntgenstruktur eines Komplexes mit Trypsin wurde bereits erwähnt (Chiche *et al.* 1989), allerdings wurden keine Koordinaten für diesen Komplex in der PDB deponiert. Das Protein wurde bisher entweder aus der Spritzgurke direkt gewonnen (Favel *et al.* 1989) oder chemisch synthetisiert, wobei die *in vitro* Faltung im Gegensatz zu den ω-Conotoxinen vollständig und spontan durch Oxidation mit Luftsauerstoff erfolgte (Le-Nguyen *et al.* 1989; Wentzel *et al.* 1999).

Für EETI-II konnte ein Intermediat im Faltungsweg weitgehend mittels 2D-NMR aufgeklärt werden. Zuerst werden die Disulfidbrücken zwischen den Cysteinresten #9/#21 und #15/#27 geschlossen, zuletzt erfolgt die Verknüpfung der Cysteinreste #2 und #19 (Le-Nguyen *et al.* 1993). Weitere kinetische Untersuchungen zeigten, dass der von den Aminosäureresten #22 bis #25 gebildete *b-turn* eine wichtige Rolle bei der Faltungskinetik von EETI-II spielt (Wentzel *et al.* 1999). Auch bei amino-terminaler Verkürzung der Polypeptidkette auf ein 23 bzw. 21 Aminosäurereste umfassendes Oligopeptid lässt sich mit NMR-Mehoden eine der nativen Faltung des Inhibitors nahe Vorzugskonformation in Lösung feststellen (Heitz *et al.* 1999). EETI-II ist mit einer Komplexdissoziationskonstante von 8×10^{-11} M einer der stärksten bekannten Trypsininhibitoren.

Abb. 6: Bild der Spritzgurke *Ecballium elaterium*. Der Trypsin Inhibitor II wurde aus den Früchten isoliert, die auf dieser Photographie der blühenden Pflanze noch nicht ausgebildet sind. Photo: Jim Manhart[2]

[2]Jim Manhart, Department of Biology, Texas A&M University, College Station, TX 77843, (979) 845-3356

Die Tatsache, dass von den oben beschriebenen Mikroproteinen bisher nur NMR-Strukturen und Röntgenstrukturdaten aus Komplexen zur Verfügung stehen, kann damit erklärt werden, dass Verbindungen dieser Art bis vor kurzem noch ein Problem für die Röntgenstrukturanalyse darstellten: Aufgrund der hier nur ansatzweise vorhandenen Ausbildung von Sekundär-Strukturelementen wird die Kristallisation durch die flexibilitätsbedingte Inhomogenität der Moleküle sowie ihre hohe Löslichkeit in polaren Lösungsmitteln erschwert. Daher müssen klassische Kristallisationsmethoden häufig variiert werden, um von mikro-kristallinem Präzipitat zu Kristallen zu kommen, die auf einer starken Röntgenquelle vermessen werden können. Die Methoden zur Strukturlösung von kleineren anorganischen und organischen Verbindungen sind ungeeignet, da die beschriebenen Polypeptide zu viele Atome beinhalten. Methoden der Proteinkristallographie wie z. B. *molecular replacement* lassen sich hingegen nur eingeschränkt anwenden, da aufgrund der weniger ausgeprägten Sekundärstruktur die molekulare Flexibilität der Mikroproteine zu unterschiedlichen Konformationen der Verbindungen im Kristall führt. Hinzu kommt, dass in einer Patterson-Funktion der Unterschied zwischen inter- und intramolekularen Abstandsvektoren viel geringer als bei größeren Proteinen ausfällt, woran die Rotations- bzw. Translationssuche mittels Modellfragment oft scheitert. Andererseits ist bei Verbindungen in der Größenklasse der Conotoxine, die im allgemeinen in einer relativ dichten Packung mit geringem Lösungsmittelanteil kristallisieren, die Erzeugung isomorpher Schweratomderivate häufig unmöglich, da die Molekülpackung dadurch gestört würde.

Mit der Entwicklung der sogenannten *dual space* Methoden ist es in den letzten 3-4 Jahren gelungen, Röntgenstrukturen ähnlich großer Substanzen zu den hier beschriebenen aus atomar aufgelösten Datensätzen zu lösen. Der den *dual space* Methoden zu Grunde liegende *shake-and-bake* Algorithmus (Miller *et al.* 1993) arbeitet bei der Strukturlösung mit iterativen Wechseln zwischen reziprokem und realem Raum. Am Anfang der Strukturlösung werden Phasen für einen Satz von Atomen erzeugt, für den bestimmte Vorgaben, wie z. B Mindestabstände zueinander, gemacht werden können. Diese Start-Phasen werden im reziproken Raum mit Direkten Methoden optimiert (*shaking*), mit diesen Phasen wird dann eine Dichtekarte im realen Raum berechnet und anschließend modifiziert (*baking*). Aus ausgewählten Maxima dieser Elektronendichte werden neue Phasen für den reziproken Raum erzeugt, die dann wiederum mit Direkten Methoden weiter verfeinert werden. Durch viele aufeinander

folgende *shake-and-bake* Zyklen kann u. U. eine Struktur *ab initio* gelöst werden. Eine wesentliche Voraussetzung für den Erfolg der *shake-and-bake* Methode ist die atomare Auflösung der gemessenen Daten. In einigen Fällen konnten auch Strukturen aus Daten, deren Auflösung niedriger als 1.2 Å war, gelöst werden.

Ein weitere Begrenzung finden die *dual space* Methoden in der Größe der zu lösenden Struktur. Zum Zeitpunkt dieser Arbeit ist die größte mit dem *shake and bake* Verfahren gelöste, nur Leichtatome enthaltende Struktur Lysozym mit 1001 Nicht-Wasserstoff-Atomen. Schweratome wie z. B. nativ eingebautes Eisen in Cytochromen können das Auflösungslimit lockern, gleiches gilt für eine hohe Schwefelatom-Dichte: Die Struktur des 402 Atome enthaltenden cystinreichen Proteins Hirustasin konnte beispielsweise bei einer Auflösung von 1.55 Å (Usón *et al.* 1999) bestimmt werden. Gegenwärtig ist der *shake-and-bake* Algorithmus in den Programmen SnB (Miller *et al.* 1993) und SHELXD (Sheldrick 1997) implementiert, wobei sich beide Programme in wichtigen methodischen Details voneinander unterscheiden. Mit dem Programm SnB konnten beispielsweise die Kristallstrukturen von α-Conotoxinen bestimmt werden: Das 13 Aminosäuren große α-Conotoxin GI aus *Conus geographus* wurde mit Daten bis 1.2 Å Auflösung *ab initio* gelöst (Guddat *et al.* 1996). Mit SHELXD wurde neben dem oben erwähnten Hirustasin unter anderem die Struktur von Vancomycin (Schäfer *et al.* 1996) gelöst. Es ist anzunehmen, dass durch die beschriebenen Programme neue Substanzklassen von Polypeptiden und kleineren Proteinen für die Röntgenstrukturanalyse erschlossen werden können.

Der *shake-and-bake* Algorithmus kann auch zur Lösung von Substrukturen aus Schweratomen oder anomal streuenden Atomen eingesetzt werde. Diese Anwendung der *dual space* Methoden ist bei SAD-Experimenten sehr erfolgreich (Dauter & Dauter 2001), eine Zusammenstellung aller mit SnB gelösten Substrukturen findet sich unter der URL: http://www.hwi.buffalo.edu/SnB/SnBSuccesses.htm .

2 Aufgabenstellung

Im Rahmen dieses Promotionsvorhabens sollten Inhibitor-Cystinknoten-Proteine mit Methoden der Röntgenstrukturanalyse charakterisiert werden. Hierfür sollte zunächst die heterologe Genexpression und präparative Aufreinigung von EETI-II zu kristallographischen Zwecken etabliert werden. Die Kristallstruktur des *Ecballium elaterium* Trypsin Inhibitors-II als Modellverbindung für Cystinknoten-Proteine sollte frei und im Komplex mit Trypsin bestimmt werden. Darüber hinaus sollten einige Varianten von EETI-II röntgenkristallographisch untersucht werden. Auf der Grundlage der aus der Kristallisation und Strukturanalyse von EETI gewonnen Erfahrungen sollten erste Experimente zur Bestimmung der Strukturen verwandter Mikroproteine durchgeführt werden.

3 Materialien und Methoden

3.1 Biologische Materialien

Bakterienstamm:

***E. coli* W3110** (Preuss *et al.* 1990)

Genotyp: F^- Lambda-, IN(rrnD-rrnE)1

Expressionsvektor:

PLZPWB-ETI-II

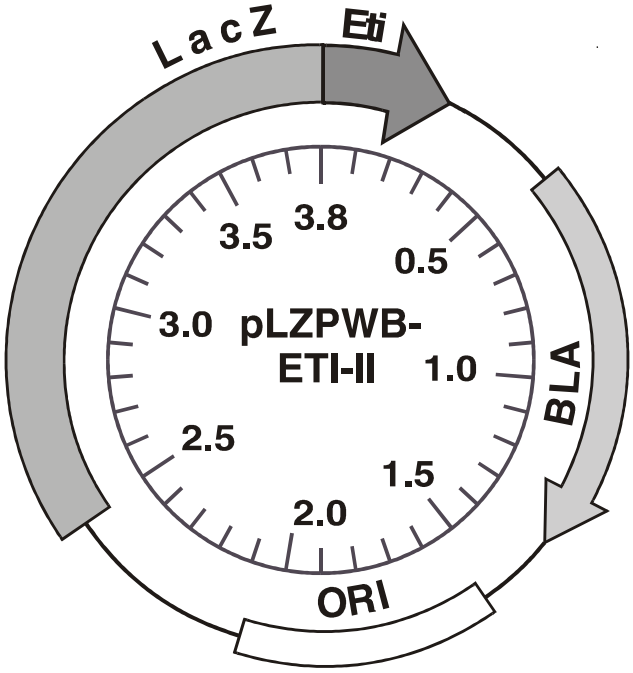


Abb. 7: Schematische genetische Karte des Vektors pLZPWB ETI-II. Die Pfeilrichtungen zeigen die funktionellen Orientierungen der einzelnen Elemente an. BLA: codierende Sequenz für b-Lactamase, LacZ: codierende Sequenz für das modifizierte b-Galaktosidase-Fragment des Fusionsproteins, Eti: codierende *eeti*-Sequenz für das EETI-II-Fragment des Fusionsproteins, ORI: Replikationsursprung. Die Skala bezeichnet die Position der Gene in der Einheit von Kilobasen.

Molekulargewichtsmarker für Proteine:

Zur Molekulargewichtsbestimmung von Proteinen mittels PAGE-Gelen wurde der „*low-range*”-Standard (BioRad Laboratories GmbH, München) verwendet:

Protein	Herkunft	Rel. Molekulargewicht
Phosphorylase b	Kaninchenmuskel	97400
Rinderserum Albumin	Rinderserum	66200
Ovalbumin	Hühnereiweiß	42700
Kohlensäure-Anhydrase	Rind	31000
Trypsin-Inhibitor	Sojabohne	21500
Lysozym	Hühnereiweiß	14400

Trypsin:

Aus dem Schweinepankreas (E.C. 3.4.21.4) von Sigma, Deisenhofen. Katalognr. T7418

Terriffic Broth (Nährmedium): Zur Proteinproduktion

24 g Hefe-Extrakt, 12 g Trypton, 4 ml Glycerol, NaCl, H_2O ad 900 ml, nach Autoklavieren: Zugabe 2.31 g KH_2PO_4 und 12.54 g K_2HPO_4 in 100 ml H_2O (steril).

dYT-Medium (Nährmedium): Zur Anzucht der Vorkultur

10 g Hefe Extrakt, 16 g Trypton, 5 g NaCl, H_2O ad 1 l, autoklaviert.

Pufferlösungen und sonstige Materialien:

Ampicillin-Stammlösung

100 mg/ml Ampicillin (Na-Salz) in Wasser, steril filtriert.

APS-Stammlösung

10% (w/v) Ammoniumperoxodisulfat, frisch angesetzt.

Coomassie-Färbelösung

2.5 g Coomassie Brilliant blue, 500 ml Methanol, 500 ml 20% (v/v) Essigsäure

Coomassie-Entfärbelösung

500 ml Methanol, 500 ml 20% (v/v) Essigsäure

Screening* Lösungen für das *sparse matrix sampling

Crystal screens I und II: Hampton Research Inc., CA, USA

IB-Waschpuffer

50 mM Tris/HCl pH 7.5, 0.5 % Triton X

IPTG-Stammlösung

1 M Isopropyl-1-thio-β-D-galactopyranosid in Wasser, steril filtriert.

***IMAC*-Puffer**

100 mM NaH_2PO_4/Na_2HPO_4 pH 8.0, 100 mM NaCl, 0.5 mM DTT.

dn*IMAC*-Puffer

100 mM NaH_2PO_4/Na2HPO4 pH 8.0, 100 mM NaCl, 8M Harnstoff, 0.5 mM DTT.

Laemmli-Probenpuffer

62.5 mM Tris/HCl pH 6.8, 10 % (v/v) Glycerol, 70 mM SDS, 5 % (v/v) β-Mercaptoethanol, 0.0025 % (w/v) Bromphenolblau.

Alle Chemikalien, die nicht explizit erwähnt sind, wurden von der SIGMA-ALDRICH Chemie GmbH, Taufkirchen bezogen.

3.2 Molekularbiologische Methoden

3.2.1 Stammhaltung und Vermehrung von E. coli Stämmen

Zur längerfristigen Aufbewahrung von plasmidtragenden *E. coli*-Stämmen wurden Glycerolkulturen angelegt. 1.5 ml autoklavierten Glycerols wurden hierfür mit 1 ml einer Übernachtkultur des betreffenden Bakterienstammes vermischt. Die Glycerolkulturen wurden bei –20 °C gelagert. Der in dieser Arbeit verwendete *E.coli* Stamm W3110 wurde bei 37 °C in dYT-Medium vermehrt. Im Falle von plasmidcodierter Resistenz wurden die Medien mit dem Antibiotikum Ampicillin komplettiert. Zur Proteinproduktion wurden 50 ml dYT-Medium mit Zellen aus

Glycerinkulturen beimpft. Volumina von 1 Liter wurden mit 25 ml einer über Nacht gewachsenen Vorkultur beimpft. Alle Kulturen wurden bei 150 Upm geschüttelt. Die Zelldichte der Kulturen wurde mittels Messung der optischen Dichte bei 600 nm im Photometer bestimmt. Im Bedarfsfall wurde mit sterilem Medium soweit verdünnt, bis die OD_{600} unter 1 lag.

3.2.2 Vorbereitung von Geräten und Lösungen

Alle in Hitze stabilen Geräte wurden zur Sterilisation für 20 Minuten auf 121 °C erhitzt. Hitzestabile Lösungen wurden für 20 Minuten bei 121 °C autoklaviert. In Hitze unbeständige Geräte wurden mit 70 % (v/v) Ethanol gereinigt und getrocknet. Hitzelabile Bestandteile von Lösungen wurden als konzentrierte Stammlösungen angesetzt und vor der Zugabe zu bereits sterilen Lösungen durch einen Membranfilter der Porengröße 0.2 µm steril filtriert.

3.2.3 Transformation von E. coli nach der Elektroporationsmethode

Zur Transformation von *E. coli* Zellen nach der Elektroporationsmethode (Dower *et al.* 1988) wurde eine 50 ml Kultur des Stammes *E. coli* K12 W3110 bis zu einer OD_{600} zwischen 0.4 und 0.8 angezogen. Die Zellen wurden durch Zentrifugation (4000 x g, 10 min, 4 °C) geerntet und in 25 ml eiskaltem, sterilem Wasser resuspendiert und erneut abzentrifugiert. Nach einmaliger Wiederholung dieses Schrittes wurde das Zellpellet in 10 ml eiskalter Glycerollösung (10 % (v/v)) aufgenommen, nach Abzentrifugieren (4000 x g, 10 min, 4 °C) wurden die Zellen in 200 µl eiskalter Glycerollösung resuspendiert. Zur Transformation wurden 40 µl Zellsuspension mit einer von der gewünschten Transformantenrate abhängigen Menge DNA versetzt, für 5 min auf Eis inkubiert und dann in eine vorgekühlte Elektroporationsküvette pipettiert. Für die Elektroporation wurde ein Genepulser (BioRad) bei folgenden Einstellungen verwendet: Stromstärke: 200 mA, Spannung: 2.5 kV, Feldstärke : 6250 V/cm, Zeitkonstante: 4.5 ms und Temperatur: Raumtemperatur.

Unmittelbar nach dem Potentialpuls wurde 1 ml steriles dYT Medium in die Küvette pipettiert. Die suspendierten Zellen wurden 1 h bei 37 °C auf einem Heizblock inkubiert und anschließend zum Beimpfen einer Flüssigkultur oder zum Ausstreichen auf Agar verwendet.

3.2.4 Spaltung von Proteinen an Methionresten mit Bromzyan

Proteine und Polypeptide lassen sich in einer hochspezifischen und nahezu quantitativ verlaufenden Reaktion mit Bromzyan spalten. Die Reaktion wurde in 70 % (v/v) Ameisensäure mit einem 100-fachen molaren Überschuss an Bromzyan zu Methionin unter Lichtausschluss durchgeführt. Zur Berechnung der Molarität des Fusionsproteins in Lösung wurde das Trockengewicht der Einschlusskörper mit einem Viertel ihres Nassgewichts berechnet. Das Reaktionsvolumen betrug 10 ml/g Einschlusskörper. Zum Ausschluss von Luftsauerstoff wurde die Lösung mit Argon überschichtet. Als Nebenreaktion kann eine partielle saure Hydrolyse von Aspartat-Prolin-Bindungen stattfinden. In Abb. 8 ist der Reaktionsverlauf schematisch skizziert.

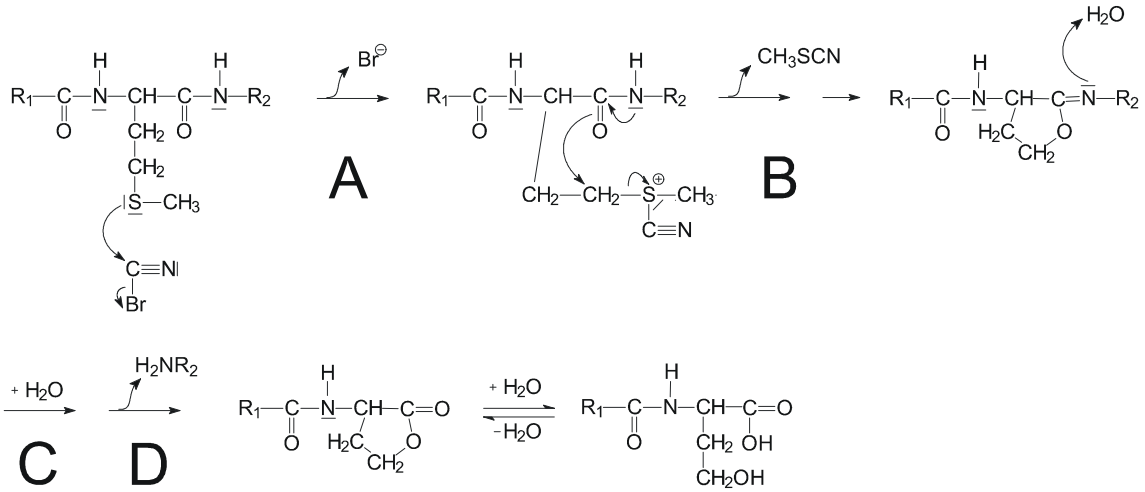


Abb. 8: Bromzyanspaltung. Der erste Reaktionsschritt (A) ist ein nukleophiler Angriff des Methioninschwefels am Kohlenstoff von Bromzyan, wodurch es zur Bildung eines Sulfonium-Ions kommt. Unter Freisetzung von Methylthiocyanat (B) wird ein Imminolacton gebildet. Die Hydrolyse (C) des Imminolactons führt unter Spaltung der Peptidbindung (D) zu einem Homoserin-Lacton, das in wässriger Lösung mit einem Homoserin im Gleichgewicht steht.

3.2.5 Denaturierende Polyacrylamidgelelektrophorese

Proteingemische wurden zu analytischen Zwecken über SDS-haltige Polyacrylamidgele unter denaturierenden Bedingungen aufgetrennt. Die Dimension der Gele betrug 100 x 60 x 0.75 mm^3. SDS wirkt als Detergenz, das sekundäre und tertiäre Strukturelemente der Proteine auflöst. Im Rahmen dieser Arbeit wurden ausschließlich Tris-Tricin Gele verwendet (Schägger & von Jagow 1987), die für die Trennung kleiner Proteine wie EETI-II entwickelt wurden. Es wurden immer fünf oder zehn Gele auf einmal gegossen (Gießstand: Hoefer Scientific). Die Zusammensetzungen für Sammelgel und Trenngel waren die folgenden:

Lösung	**Trenngel**	**Sammelgel**
30% (w/v) Acrylamid, 0.8 % Bisacrylamid (Protogel®)	20 ml	1.6 ml
3M Tris/HCl, 0.3 % SDS, pH 8.45	10 ml	3.1 ml
H_2O (bidest.)	-	7.6 ml
APS-Stammlösung	100 µl	100 µl
TEMED	10 µl	10 µl

Der Laufpuffer für die Kathodenkammer war 0.1 M Tris, 0.1 M Tricine, 0.1 % SDS, die Anodenkammer wurde mit 0.2 M Tris/HCl pH 8.9 befüllt. Als Längenstandard wurde *low-range* Marker (BioRad) verwendet. Als Probenauftragspuffer wurde Laemmli-Puffer verwendet. Die Elektrophorese wurde bei einer maximalen Leistung von 6 W durchgeführt. Zum Färben wurde das Gel in einem geeigneten Behälter unter leichtem Schütteln für 5 Minuten in 20 ml frisch angesetzter Coomassie-Färbelösung inkubiert. Die Entfärbung erfolgte über 30 min unter leichtem Schütteln in 50 ml Coomassie-Entfärbelösung.

3.2.6 Reversed Phase High Performance/Pressure Liquid Chromatography

Das Grundprinzip der Chromatographie besteht in der Auftrennung von Stoffgemischen, die in einer mobilen Phase (Lösungsmittel oder Gas) über eine stationäre Phase (Säulenmatrix) transportiert werden, wobei die einzelnen Stoffe des Gemisches in Folge unterschiedlich starker Interaktion mit der stationären Phase voneinander separiert werden können. Polypeptidgemische werden in der Regel über hydrophobe Säulenmaterialien als stationärer Phase, sogenannte „Umkehrphasen“ (engl. *reversed phase*) aufgetrennt. Das Säulenmaterial besteht häufig aus Kieselgel als Trägermaterial, das mit Alkylresten der Länge C_2 bis C_{18} (*n*-Octadekyl) besetzt ist. Die einzelnen Partikel haben bei analytischen und halbpräparativen Säulen einen Durchmesser von 5 µm, die Porengrößen liegen in der Größenordnung von 10-30 nm. Die mobile Phase setzt sich aus Wasser, einem organischen, hydrophoben Lösungsmittel und einem Ionenpaar-Reagenz zusammen. Das Ionenpaar-Reagenz bindet die geladenen Gruppen des Proteins und verändert damit dessen Hydrophobizität. Üblicherweise wird Trifluoressigsäure in einer Konzentration von 0.1% (v/v) zu diesem Zweck eingesetzt. Der Volumenanteil des organischen Lösungsmittels wird während des chromatographischen Laufs in einem kontinuierlichen Gradienten erhöht, wobei das hydrophobe Lösungsmittel in zunehmendem Maße mit den Peptiden um die lipophilen Bindungsplätze der stationären Phase konkurriert.

Als organisches Lösungsmittel wird sehr oft Acetonitril eingesetzt, das eine geringe Viskosität und gute UV-Durchlässigkeit besitzt. Die RP-Chromatographie wird aufgrund der erforderlichen hohen Trennleistung unter hohem Lösungsmitteldruck als *HPLC* (High Performance/Pressure Liquid Chromatography) durchgeführt. Zwei Pumpen, die den für diese Chromatographie notwendigen Druck von 100 bis 250 bar aufbauen, fördern den für den entsprechenden Zeitpunkt des Gradienten jeweils erforderlichen Volumenanteil beider Lösungsmittel in eine Mischkammer, von der aus die mobile Phase über einen Injektor zur Auftragung des Analytgemisches auf die Säule gelangt. Die Detektion des Elutionsprofils erfolgt über UV-Absorptionsmessung bei 217 nm. Im Falle aromatischer Reste im Peptid kann eine zusätzliche Detektion bei 280 nm erfolgen. Das Elutionsvolumen kann mit einem Fraktionssammler aufgefangen werden. Alternativ dazu können die mit einem *peak* korrelierten Volumina gezielt von Hand gesammelt werden.

3.2.7 Affinitätschromatographie über immobilisierte Metallionen (IMAC)

Bei der Proteinaufreinigung über *IMAC (Immobilized Metal ion Affinity Chromatography)* wird die Komplexbildung zwischen einem auf der Säulenmatrix immobilisierten Metallkation und einem dem Protein angehängten Oligo-Histidinabschnitt als Interaktionsprinzip zugrunde gelegt (Porath *et al.* 1983). Hierzu wird die Polypeptidkette des Proteins auf genetischer Ebene um vier bis sechs terminale Histidinreste (engl. *His-tag*) erweitert. Die Aufreinigung erfolgt über eine Säulenmatrix, die chelatbildende Gruppen trägt. Die ringständigen Stickstoffatome der Histidine des Zielproteins interagieren über die freien Elektronenpaare mit den von der Säulenmatrix komplexierten Metallatomen, was die Durchflussgeschwindigkeit für dieses Protein reduziert. In dieser Arbeit wurde *chelating sepharose* (Pharmacia/LKB GmbH, Freiburg) eingesetzt. Eine Säule mit einem Volumen von 30 ml (Höhe 20 cm) wurde mit 8-10 ml *chelating sepharose* bepackt und mit Nickelionen beladen, indem 20 ml einer 100 mM Lösung von $NiCl_2$ über die Säule gegeben wurde. Nach dem Waschen mit Wasser deutete die grüne Färbung der Säulenmatrix darauf hin, dass die Nickelionen auf der Matrix immobilisiert waren. Nach dem Äquilibrieren der Säule mit 20 ml dn-*IMAC*-Puffer wurde das aufzureinigende Protein in einem Lösungsvolumen von 20-30 ml auf die Säule gegeben, nach Waschen mit *IMAC*-Puffer wurde das aufzureinigende Protein in einem Stufengradienten ansteigender Imidazolkonzentration in *IMAC*-Puffer von der Säule eluiert. Imidazol ist eine zum Histidin strukturanaloge Verbindung, die mit dem *His-tag* um die Nickelkationen konkurriert. Die einzelnen Fraktionen aus der *IMAC* wurden mittels *PAGE*-Gel analysiert. Die *IMAC*-Säule wurde durch Waschen mit 30 ml EDTA-Lösung (0.5 M) und nachfolgender Neubeladung mit Nickel-Ionen regeneriert.

3.2.8 Bestimmung der Ausbeute an EETI-II

Da natives EETI-II nur einen Phenylalaninrest und keine weiteren Chromophore enthält, ist seine Konzentrationsbestimmung in Lösung über die Absorption bei 280 nm nicht mit ausreichender Präzision durchführbar. Die gravimetrische Bestimmung von EETI-II nach der Lyophilisation ist ebenfalls mit Nachteilen behaftet, da im Bereich der

erzielten Ausbeute von 1-3 mg Inhibitor pro Liter Bakterienkultur die Wägung von systematischen Fehlern behaftet ist. Darüber hinaus ist das lyophilisierte Polypeptid meist elektrostatisch geladen und kann aufgrund seiner watteartigen Konsistenz sehr leicht verweht werden. Für die Konzentrationsbestimmung unter Laborbedingungen wurde daher einmalig eine Korrelation zwischen der Fläche des EETI-II-*peaks* aus dem *HPLC*-Verlauf mit der danach eingewogenen Menge an Polypeptid vorgenommen. Im Folgenden wurde die Konzentration der EETI-II-Lösung zu Zwecken der Reproduzierbarkeit auf die Flächen der *peaks* aus dem *HPLC*-Diagramm bezogen eingestellt.

3.3 Kristallographische Methoden

3.3.1 Umgang mit Proteinlösungen zu kristallographischen Zwecken

Alle Vorratslösungen von Proteinen zur Kristallzucht wurden bei 4 °C aufbewahrt. Nach dem Ansetzen wurden die Lösungen für 5 Minuten bei 12000 UpM zentrifugiert, der Überstand wurde in ein neues Reaktionsgefäß transferiert. Nach längerer Standzeit wurden die Lösungen unter gleichen Bedingungen zentrifugiert, um eventuelle Präzipitate zu entfernen.

3.3.2 Kristallisation von Proteinen

Bei der Kristallzucht wurde die Gas-Diffusions-Methode zwischen einer Reservoirlösung und einem „hängenden Tropfen“ angewendet (Ducruix & Griegé 1992, McPherson 1998). Die Kristallisation des Makromoleküls soll bei diesem Verfahren über die langsame und stufenlose Erhöhung der Konzentrationen aller in Lösung vorliegenden Substanzen einschließlich des Proteins erreicht werden. Dies wird über den Ausgleich der chemischen Potentiale zwischen einem Tropfen der Proteinlösung und einem viel größeren Volumen einer „Reservoir-Lösung“ über Gasphasendiffusion herbeigeführt. In der Praxis wird hierzu ein Tropfen der Proteinvorratslösung mit einem zweiten Tropfen der Reservoirlösung in einem bestimmten Volumenverhältnis vermischt. Die Proteinvorratslösung sollte neben dem Makromolekül nur geringe

Konzentrationen an Salz und Puffer enthalten. Die Konzentration an Protein ist ein Parameter, der im weiteren Verlauf der Kristallisationsversuche verändert werden kann. Eine übliche Ausgangskonzentration ist 10 mg/ml. Die Reservoir-Lösung kann beispielsweise Puffersubstanzen, Präzipitanden wie Polyethylenglykole, Salze, Additive wie Spermin, organische Lösungsmittel oder andere Bestandteile enthalten.

In der Praxis werden die Vertiefungen in Gewebekulturplatten als Reservoir-Behälter eingesetzt. Um das Diffusions-System gegen die Umgebung abzudichten, werden die Ränder der zylindrischen Vertiefungen in den Gewebekulturplatten mit Silikonfett bestrichen. Der „hängende Tropfen“ wird auf einem zur Erhöhung der Oberflächenspannung silikonisierten Deckgläschen gemischt, das daraufhin vorsichtig gewendet und über die Reservoirlösung gelegt wird. Nach leichtem Andrücken schließt das Deckglas luftdicht mit dem Rand der zylindrischen Vertiefung ab, wodurch ein geschlossenes System entsteht.

Am Anfang der Etablierung von Kristallisationsbedingungen steht bei unbekannten Proteinen meist ein ausgedehntes *screening*, wobei bis zu mehrere 100 unterschiedliche Lösungen auf ihre kristallisationsfördernde Wirkung hin getestet werden. Im Rahmen dieser Arbeit wurden für diese als *sparse matrix sampling* (Jarmila & Kim, 1991) bezeichnete Methode die käuflichen *crystal screens* I und II der Firma Hampton Research eingesetzt. Auf der Grundlage der Ergebnisse aus diesen *screens* können Parameter wie die Zusammensetzung der Reservoirlösung, ihre individuellen Stoffkonzentrationen, die Kristallisationstemperatur oder die Proteinkonzentration systematisch verändert werden. Diese Verfeinerung von Bedingungen kann im Idealfall zu Proteinkristallen führen, die für die Röntgenstrukturanalyse geeignet sind.

Eine Möglichkeit, um aus Bedingungen für kleine oder unregelmäßig gewachsene Kristalle größere zu generieren, ist das sogenannte *seeding*. Hierfür werden die Kristallisationsbedingungen so variiert, dass die Endkonzentrationen im hängenden Tropfen nach dem Diffusionsausgleich nicht mehr zur Kristallkeimbildung führen. Durch Überimpfen eines kleinen Kristalls (*makroseeding*) oder submikroskopischer Keime (*microseeding*) kann unter Umständen das Wachstum von Kristallen ausgelöst werden, die besser für die Röngenstrukturanalyse geeignet sind. Zwischen den ersten Bedingungen und den messbaren Kristallen können mehrere Optimierungsschritte liegen.

3.3.3 Kristallmontage

Zur Durchführung von Röntgenbeugungsexperimenten müssen die Einkristalle so auf dem Messgerät montiert werden, dass sie auf dem Goniometerkopf fest fixiert und im Strahlengang zentriert sind. Hierbei sollte nach Möglichkeit nur der zu messende Kristall im Strahlengang liegen, da andere, auch amorphe Festkörper oder Flüssigkeiten zu einer Erhöhung des Hintergrundrauschens auf dem Detektor führen. Während der Vermessung und der damit verbundenen Veränderung der Goniometer-Winkelgeometrie darf der Kristall seine relative Position zum Goniometerkopf nicht verändern. Für die Vermessung der meist sehr empfindlichen Proteinkristalle empfiehlt sich die Montage in einer Kunstfaser-Schlaufe (*loop*) unter Kryo-Bedingungen (Garman 1999). Wichtige Vorteile dieser Methode ergeben sich daraus, dass der Kristall während der gesamten Messung bei 100 K gefroren ist: Unter diesen Bedingungen ist die thermische Bewegung der Atome im Kristall geringer, d. h. ihre Positionen lassen sich mit geringeren Standardabweichungen als bei Rautemperatur bestimmen. Darüber hinaus treten sekundäre Strahlenschäden als Folge von röntgeninduzierter Radikalbildung gegenüber Raumtemperatur-Messungen stark verlangsamt auf. Gefrorene Kristalle können zuerst im eigenen Labor, beispielsweise auf einem Goniometer mit einer Drehanode, getestet werden und dann zur intensiveren Untersuchung in gefrorenem Zustand in einem Trockendewar zu einem Synchrotron transferiert werden. In der Praxis wird der Kristall hierzu aus einer Lösung gezüchtet, die unter plötzlichem Gefrieren in flüssigem Stickstoff amorph erstarrt. Falls dies nicht möglich ist, kann der Kristall vor dem Montieren einige Sekunden bis Stunden in einem Kryo-Puffer inkubiert werden. Der Kryo-Puffer sollte so gewählt werden, dass sich die Mosaizität des Kristalls während des Schockgefrierens nicht wesentlich erhöht. Typische Kryo-Protektanden, deren Zugabe einen Puffer in einen Kryo-Puffer überführt, sind beispielsweise Glycerol, Isopropanol, MPD oder Malonat.

Zur Montage wird der Kristall mit einer Schlaufe, deren Durchmesser idealer Weise ca. 0.1 mm größer als der Kristall sein sollte, aus dieser Lösung entnommen und umgehend in flüssigen Stickstoff gefroren. Die Schlaufe ist hierbei an einem Stahlstift befestigt, der in eine Haltevorrichtung, den sogenannten *top hat* aus Nickel eingeschraubt ist. Da Nickel ferromagnetisch ist, kann diese Konstruktion mittel eines auf dem Goniometerkopf befestigten Magneten fixiert werden. Um den Kristall im gefrorenen

Zustand in den Kaltgasstrom des Messgerätes zu bringen, hat sich die Verwendung einer gebogenen Metallschiene, des sogenannten „*arc*“ bewährt (Abb. 9).

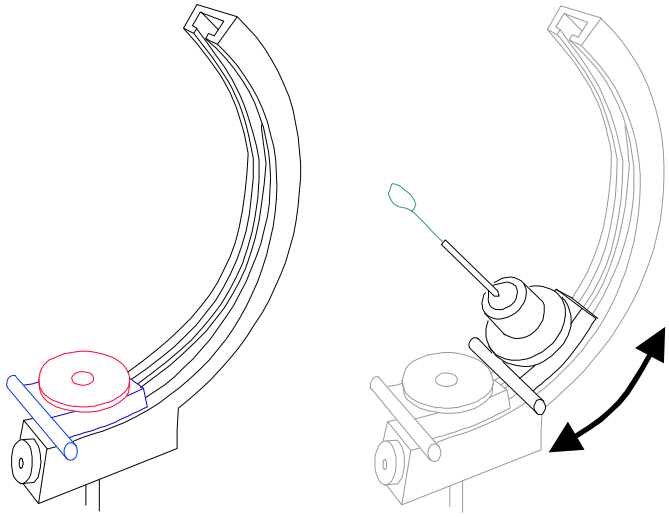


Abb. 9: Links: Eine wichtige Hilfe bei der Montage gefrorener Kristalle ist eine als *arc* bezeichnete, gebogene Aluminum-Schiene (schwarz), auf der ein Schlitten (blau) gegen leichten Reibungswiderstand verschoben werden kann. Auf diesem Schlitten ist eine magnetische Platte (rot) montiert, auf der der *top hat* fixiert werden kann. Rechts: Der *loop* (grün) ist in einem Stahlstift eingeklebt, der im *top hat* mit einer Schraube befestigt wird.

Auf diesem *arc* lässt sich ein Magnet auf einem Schlitten gegen leichten Reibungswiderstand bewegen. Zum Montieren wird der Schlitten über den Kaltgasstrom geschoben. Der im *loop* gefrorene Kristall wird direkt im flüssigen Stickstoff in ein Transportgefäß getaucht, auf dessen Rändern die Basis des *top hat* aufliegt. Mit einer speziellen Pinzette wird dieses Transportgefäß zügig so zum Magneten orientiert, dass die Basis des *top hats* von diesem angezogen wird. Durch schnelles Wegziehen des Behälters nach unten kommt der Kristall aus dem flüssigen Stickstoff direkt in den Kaltgasstrom und kann durch Verschieben des Schlittens so orientiert werden, dass der Metallstift auf der Phi-Achse liegt. Nach dem Zentrieren kann der *arc* entfernt werden.

Wenn keine geeigneten Kryo-Bedingungen für einen Kristall gefunden werden können, muss die Datensammlung bei Raumtemperatur vorgenommen werden. Da für Proteinkristalle ein hoher Wasseranteil charakteristisch ist, müssen sie zum Schutz vor Austrocknung in eine Quarzglaskapillare montiert werden, deren Enden mit Wachs verschlossen werden. Hierzu wird der Kristall zunächst zusammen mit Mutterlauge in

eine Kapillare aus Quarzglas gesaugt. Da Glas Röntgenlicht in geringem Ausmaß unspezifisch streut, sollte der Durchmesser der Kapillarwand zur Unterdrückung einer zu hohen Hintergrundstrahlung möglichst dünn gewählt werden. Während der Messung darf sich der Kristall nicht relativ zur Kapillare und zum Goniometerkopf bewegen. Daher wird die Mutterlauge mit einem haarfein geschnittenen Filterpapierstreifen soweit abgesaugt, dass nur noch an den beiden Enden der Kapillare einige kleine Tropfen Pufferlösung übrig bleiben, die eine Sättigung der Luftfeuchtigkeit in der Glaskapillare sicherstellen. Der Kristall sollte, von wenig Lösungsmittel umgeben, an der Kapillarwand anliegen. Die versiegelte Kapillare wird auf einen PVC-Stift geklebt, der auf dem Goniometerkopf zur Messung fixiert wird.

3.3.4 Röntgenquellen

3.3.4.1 Drehanode

Eine im Labor gut etablierte Röntgenquelle zu Zwecken der Proteinkristallographie ist die Kupfer-Drehanode. Während des Betriebs rotiert ein ringförmiger Kupferzylinder als Anode kontinuierlich unter einem auftreffenden Kathodenstrahl hindurch, was eine sehr effiziente Kühlung des Metalls erlaubt. Die im Rahmen dieser Arbeit verwendete Drehanode wurde bei einer Leistung von ca. 5 kW betrieben. Aus dem Spektrum des generierten Röntgenlichts wird nur die intensive K_α-Strahlung mit einer Wellenlänge von 1.54178 Å für Beugungsexperimente an Proteinkristallen verwendet. Um das Licht dieser Wellenlänge aus der Bremsstrahlung herauszufiltern, werden im wesentlichen zwei Systeme eingesetzt: Weit verbreitet ist die Nutzung eines Graphit-Einkristalls, dessen 002-Reflex für die entsprechende Wellenlänge durch das Kollimator-Rohr auf den zu messenden Kristall trifft. Als weiteres, in der Anschaffung wesentlich teureres System, findet eine *multi layer* Optik (z. B. von Osmic) Anwendung, die im Vergleich zu einem Kristall-Monochromator mit wesentlich geringeren Intensitätsverlusten arbeitet. Mit einem solchen System ist die Fokussierung des Strahls möglich.

3.3.4.2 Synchrotron

Die Intensität der Drehanoden-Strahlung reicht in vielen Fällen nicht aus, um höher aufgelöste Datensätze von Proteinkristallen zu messen. Hier bietet sich ein Synchrotron als Quelle bedeutend intensiverer Strahlung an (Helliwell 1992): Wenn Elektronen bzw. Positronen im Hochvakuum auf einer Kreisbahn bis nahe an die Lichtgeschwindigkeit beschleunigt werden, geben sie tangential zu Ihrer Flugbahn die sogenannte Synchrotronstrahlung ab. Das Synchrotronlicht ist polychromatisch, mit einem Monochromator lässt sich die Wellenlänge des zu Messzwecken eingesetzten Lichts variieren. Der Strahl ist um Größenordnungen schärfer (brillanter) als der einer rotierenden Anode, wodurch beispielsweise auch sehr kleine Kristalle bei geeigneten Blendeneinstellungen ohne hohen Hintergrund gemessen werden können. An einigen Teilchenbeschleunigern kann eine so hohe Strahlenintensität erreicht werden, dass es trotz der Kühlung auf 100 K zu deutlichen Strahlenschäden im Protein kommen kann (s. u.).

In der Praxis werden auf der Drehanode Messungen von hochredundanten Datensätzen bei niedriger Auflösung, z. B. von Schweratomderivaten, und Testmessungen vorgenommen. Datensätze, gegen die ein endgültiges Modell einer Struktur verfeinert werden soll, werden in der Regel am Synchrotron gesammelt.

3.3.5 Flächenzähler

Proteine kristallisieren im allgemeinen in großen Elementarzellen. Um einen vollständigen, hochaufgelösten Datensatz zu erhalten, müssen in einigen Fällen bis zu mehrere Millionen unabhängige Reflexe gemessen werden. Zur Zeit werden im wesentlichen drei Zählertypen zur Messung von Proteinkristallen eingesetzt:

3.3.5.1 CCD-Zähler

CCD-Zähler funktionieren nach dem Prinzip einer Videokamera (Gruner & Ealick 1995). Die Röntgenquanten eines Reflexes treffen hier zunächst auf eine Phosphor-Folie auf, die dadurch zur Emission von Licht im sichtbaren Bereich angeregt wird. Durch eine reduzierende Glasfaseroptik wird dieses Licht auf den dahinter liegenden

CCD-Chip weitergeleitet. Ein wichtiger Vorteile dieses Zählers liegt in seiner kurzen Auslesezeit. Aufgrund seines geringen Gewichtes kann ein CCD-Detektor auf einem beweglichen Goniometerkreis montiert werden, wodurch der Messbereich bis zu hohen Auflösungen erweitert werden kann. Ein CCD-Detektor eignet sich aufgrund seiner hohen Sensitivität sehr gut zur Messung von Intensitätsunterschieden zwischen Friedel-Paaren und von schwach streuenden Kristalle. Ein generelles Problem von CCD-Detektoren ist ein systemimmanentes, permanentes Untergrundrauschen. Durch kontinuierliche Kühlung des Chips auf Temperaturen von unter -50 °C lässt sich dieser Hintergrund deutlich reduzieren.

3.3.5.2 Image Plate (Amemiya et al. *1988)*

Der *image plate* Zähler besitzt eine viel größere aktive Fläche als ein CCD-Zähler. Die Detektion der Röntgenquanten erfolgt hier über eine mit BaFBr beschichtete Fläche, in der sich bei Anregung mit Röntgenquanten quasistabile Farbzentren bilden. Nach der Exposition wird das Bild mechanisch ausgelesen, wobei ein Photomultiplier eine photostimulierte Lumineszenz nach Laseranregung detektiert. Unmittelbar nach dem Auslesen wird das aufgenommene Bild durch Bestrahlung mit einer Halogen-Lampe gelöscht. Die Auslesezeit ist hierbei bedeutend höher als bei einem CCD-Zähler, wodurch an einem Synchrotron wertvolle Strahlzeit nicht genutzt werden kann. Im direkten Vergleich zu einem CCD-Zähler ergibt sich dadurch allerdings nicht zwangsläufig ein Nachteil, da ein CCD-Detektor aufgrund seiner geringeren Fläche mehrere Bilder aufnehmen muss, um die gleiche Anzahl an Reflexen wie das *image plate* zu messen. *Image plate* Detektoren wie der MAR345 sind so schwer, dass sie nicht auf einem beweglichen 2θ-Kreis montiert werden können. Hierdurch ist die höchste messbare Auflösung durch den minimal möglichen Abstand Zähler/Kristall und die bei der Messung verwendete Wellenlänge limitiert.

3.3.5.3 Multi-Wire-Proportional-Chamber (MWPC)

Neben diesen beiden Zählertypen finden auch noch ältere Entwicklungen Anwendung. Eine im Rahmen dieser Arbeit durchgeführte Messung wurden mit einem *multi-wire-proportional-chamber* detektiert (Lewis 1994). Hier treten die Röntgenquanten durch ein Berylliumfenster in eine mit Xenon gefüllte Kammer ein, in der drei senkrecht und

waagrecht verlaufende Drahtebenen aufgespannt sind, an denen eine Hochspannung anliegt. Die Röntgenquanten werden indirekt über die von ihnen ionisierten Gasmoleküle detektiert, deren Position durch die Entladung an den Spannungsdrähten bestimmt wird. Die Vorteile eines *MWPC* liegen in seiner guten Energiediskriminierung, was in einem sehr geringen Hintergrundrauschen resultiert. Aufgrund seiner geringen Größe kann der *MWPC* auf einem beweglichen Goniometerkreis montiert werden. Wichtige Nachteile sind die aufwendige Eichung des Zählers, die vor jeder Messung neu erfolgen muss, sowie die Größe der aktiven Fläche, die deutlich kleiner als bei einem *image plate* ist. Außerdem besitzt der MWPC nur eine relativ geringe dynamische Breite.

3.3.6 Datensammlung

Die Symmetrie der Kristallpackung steht in direkter Beziehung zur Symmetrie des reziproken Gitters. Je höher symmetrisch die Lauegruppe eines Kristalls ist, desto kleiner ist der Ausschnitt des reziproken Gitters, der für einen vollständigen Datensatz gemessen werden muss. Für eine Goniometer-Plattform, die über mehrere drehbare Kreise verfügt, lässt sich nach der Zellbestimmung eine effiziente Messstrategie festlegen (Dauter 1999): Der Kristall kann so zum Strahl orientiert werden, dass in möglichst wenigen Bildern ein Datensatz mit der beabsichtigten Vollständigkeit und Redundanz gesammelt wird. Durch eine optimierte Strategie kann die Messzeit gerade bei höheren Lauegruppen deutlich verkürzt werden, was bei strahlenempfindlichen Kristallen oder unter dem Aspekt zeitlicher Limitationen am Synchrotron von erheblicher Bedeutung sein kann.

Die im Rahmen dieser Arbeit beschriebenen Datensammlungen wurden mit einer Ausnahme auf Messanordnungen durchgeführt, bei denen nur die ϕ-Achse frei drehbar war. Bei dieser an Synchrotron-*beamlines* üblichen Konfiguration ist es nützlich, den zu messenden Kristall in einer Orientierung im *loop* zu frieren, in der keine Kristallkante parallel zum Strahl verläuft. Hierdurch kann ausgeschlossen werden, dass für die Raumgruppen-Bestimmung wichtige Reflexe (h00, 0k0, 00l) durch die eingeschränkte Geometrie der Plattform messtechnisch unerfassbar bleiben. Bei nur einer drehbaren Achse sind die frei wählbaren Parameter zur Messzeit-Optimierung neben der Belichtungszeit der Startpunkt in ϕ und der Drehwinkel pro Aufnahme. Letzterer darf

eine von der Elementarzelle und der Mosaizität des Kristalls abhängige Obergrenze nicht überschreiten, da es ansonsten zur Überlagerung von im reziproken Gitter benachbarten Reflexen kommen kann. Ein Computerprogramm, das im Rahmen dieser Arbeit zu Planung der Messstrategie eingesetzt wurde, ist STRATEGY (Ravelli *et al.* 1997). Das Programm COSMO, das in Zukunft von der Firma BrukerAXS vertrieben werden wird, erlaubt eine umfassende Strategieberechnung für eine nahezu beliebige Anzahl an Goniometer-Kreisen.

3.3.7 Datenprozessierung

Die von einem Flächenzähler aufgenommenen Beugungsbilder zeichnen die Reflexe beim Durchtritt durch die Ewald-Kugel auf. Die Reflex-Indizierung, einhergehend mit der Bestimmung der Elementarzelle und der Orientierungsmatrix des Kristalls, wurde mit dem Programm XDISP aus DENZO (Otwinowski & Minor 1997) durchgeführt. Die dreidimensionalen Reflexe können pro Bild nur in zwei Dimensionen erfasst werden. Hieraus ergibt sich die Notwendigkeit, das reziproke Gitter aus den einzelnen Bildern zusammenzusetzen und die räumlichen Reflexprofile zu rekonstruieren. Dieser als „Integration“ bezeichnete Schritt der Datenprozessierung wurde für fast alle hier beschriebenen Messungen mit Hilfe der Programme DENZO und SCALEPACK (Otwinowski & Minor 1997) durchgeführt. Der dieser Software zu Grunde liegende Algorithmus arbeitet im sogenannten „*thick slice modus*“, hierbei wird der Drehwinkel pro Bild so weit gewählt, dass der Großteil der gemessenen Reflexe in einem Bild voll erfasst wird. Die von DENZO für jedes einzelne Bild indizierten und analysierten Reflexe werden mit Hilfe des Programms SCALEPACK zueinander skaliert und zum reziproken Gitter zusammengesetzt. Symmetrieäquivalente Reflexe können mit SCALEPACK gemittelt werden und statistische Werte zur Qualität der Daten können bereits aus der Protokolldatei von SCALEPACK entnommen werden. In dieser Arbeit wurde zu Zwecken der Mittelung symmetrieäquivalenter Reflexe sowie zur Bestimmung der statistischen Gütekriterien wie R_{int}, R_{sym} oder dem $I/s(I)$ Verhältnis der gemessenen Daten das Programm XPREP (XPREP 1997) verwendet.

Messdaten, die auf einem MWPC gesammelt wurden, wurden mit dem Programm FRAMBO (FRAMBO 1994) indiziert und mit SAINT (SAINT 1997) integriert, das einen dreidimensionalen Algorithmus zur Rekonstruktion der Reflexe implementiert

hat. SAINT arbeitet daher, im Gegensatz zu DENZO, optimal mit im „*thin slice modus*“ gemessenen Daten: Ein Reflex sollte hier idealerweise auf mehrere aufeinanderfolgende Aufnahmen verteilt sein. An die Integration mit SAINT schließt sich die Absorptionskorrektur mit dem Programm SADABS (Sheldrick 1997) an. Für die Datenreduktion wurde auch hier das Programm XPREP eingesetzt.

3.3.8 Strukturlösung

Da die Phasen der gemessenen Reflexe in der Praxis experimentell nicht direkt zu erfassen sind, muss man die gesamte Strukturinformation aus den Reflexintensitäten ableiten. Dies stellt vor allem bei der Suche nach einem ersten, zum Datensatz passenden Molekülmodell eine Schwierigkeit dar. Zur Lösung des sogenannten Phasenproblems der Röntgestrukturanalyse finden mehrere Methoden Anwendung.

3.3.8.1 Isomorpher Ersatz

Die historisch bedeutendste Methode zur Lösung des Phasenproblems ist der isomorphe Ersatz (engl. *isomorphus replacement*). Hierbei wird ein Datensatz, der von einem Kristall des nativen Proteins gemessen wurde, auf den Datensatz eines Schweratomderivats desselben Proteins skaliert. Die Elementarzelle des Schweratomderivats darf dabei nicht weiter als maximal 1 % von der des nativen Proteins abweichen, da die Molekülpackungen ansonsten nicht mehr die für einen Phasenvergleich ausreichende Isomorphie aufweisen. Nach der Skalierung kann aus der Differenz der Reflexintensitäten zwischen den beiden Datensätzen der Strukturfaktoranteil der Schweratome errechnet werden. Nach der Bestimmung der Schweratom-Positionen im Kristall können dann erste Phasen für die Gesamtstruktur berechnet werden. Wenn ein Schweratomderivat zur Phasenbestimmung ausreicht, spricht man von einem SIR-(*single isomorphus replacement*), wenn mehrere Derivate eingesetzt werden müssen, von einem MIR-(*multiple isomorphus replacement*)-Experiment (Drenth 1994). Die Beträge der Strukturfaktordifferenzen eines SIR-Experimentes liefern nur die unteren Grenzen für die möglichen Beträge der Schweratomstrukturfaktoren, für eine genauere Phasenbestimmung muss ein MIR-Experiment durchgeführt werden.

3.3.8.2 Anomale Dispersion

Der Effekt der anomalen Dispersion bestimmter schwererer Elemente in Abhängigkeit von der Wellenlänge der Röntgenstrahlung kann bei der Strukturlösung zur Phasenbestimmung ausgenutzt werden (Drenth 1994). Allgemein gilt, dass der Streufaktor eines Atoms bei gegebener Wellenlänge im wesentlichen von dessen Elektronenzahl und dem Streuwinkel relativ zum einfallenden Strahl abhängig ist. Bei einer genaueren Beschreibung des Streufaktors muss man allerdings die Beiträge bestimmter Resonanz-Absorptions-Effekte mit berücksichtigen: Der Anteil der Thomson-Streuung liefert den Hauptbeitrag zum Streufaktor f und wird mit f_0 bezeichnet. Röntgenquanten können oberhalb eines definierten Energieniveaus auch von der Elektronenumgebung eines Atoms absorbiert und ohne Energieverlust unmittelbar wieder abgegeben werden. Dieser Effekt ist von der Wellenlänge des Röntgenlichts abhängig und liefert den Beitrag f' zum Streufaktor. Die Röntgenquanten können auch, ebenfalls in Abhängigkeit von der Wellenlänge, vom Atom absorbiert und mit zeitlicher Verzögerung wieder freigegeben werden. Diese Wechselwirkung führt zu einem zu f_0 phasenverschobenen Beitrag f''. Der Streufaktor f kann folgendermaßen beschrieben werden:

$$f = f_0 + f'(l) + if''(l)$$

Im allgemeinen sind die Beiträge von f' und f'' sehr klein, bei bestimmten Elementen mit einer Röntgen-Absorptionskante können sie bei geeigneter Wellenlänge allerdings einige Prozent von f betragen. In Folge der Phasenverschiebung von f'' weisen die Friedel-Gegenteile unterschiedliche Intensitäten auf, aus deren Differenzen sich die Lage der anomal streuenden Atome in der Einheitszelle bestimmen lässt. Wenn diese Intensitätsunterschiede sehr schwach sind, müssen die Messungen mit hoher Genauigkeit, z. B. mittels einer hohen Redundanz, durchgeführt werden. Ihre derzeit häufigste Anwendung zu Zwecken der Bestimmung erster Phasen findet die anomale Dispersion bei MAD-Experimenten (s.u.).

Für ein MAD-(*multiple wavelength anomalous dispersion*) Experiment wird die Lage der anomal streuenden Atome aus mehreren Messungen bei unterschiedlichen Wellenlängen ermittelt. Da die Koordinaten der Schweratome im Kristall sehr genau bestimmt sein müssen, hat sich in der Praxis der biologische Einbau von Seleno-Methionin zum Zeitpunkt der Proteinproduktion im Organismus durchgesetzt. Der Protein produzierende Organismus kann hierzu beispielsweise mit Minimalmedium ernährt werden, dem anstatt der Aminosäure Methionin das Derivat Selenomethionin zugesetzt wird. Diese Substitution hat in den meisten Fällen keinen wesentlichen Einfluss auf die Proteinstruktur, kann aber aufgrund der Selenatome zur Lösung des Phasenproblems ausgenutzt werden. Selen hat eine Absorptionskante bei einer Wellenlänge von 0.97945 Å (12657.80 eV). Üblicherweise werden drei Datensätze bei Wellenlängen gemessen, die als *remote*-(entfernt von der Absorptionskante), *inflection*-(bei minimalem *f'*) und als *peak*-(bei maximalem *f''*) Bereich bezeichnet werden. Bei einer guten Datenqualität ist die natürliche Häufigkeit von Methioninresten in Proteinen, die bei dieser Methode zu etwa einem Selenatom pro 50 Aminosäuren führt, für die Phasenbestimmung ausreichend. Bei einer zu geringen Selendichte wird das anomale Signal zu schwach, um ausreichende Phaseninformation ableiten zu können. Der Vorteil von MAD-Experimenten ist, dass der Einbau von Selenomethionin relativ gut etabliert ist und die Software zur Phasenbestimmung zur Verfügung steht. Ein Nachteil ist, dass MAD-Experimente sehr viel Zeit in Anspruch nehmen und empfindliche Kristalle daher nicht vollständig vermessen werden können. Daten aus MAD-Messungen können mit einem nativen Datensatz des gleichen Proteins zu einem MIRAS-Experiment kombiniert werden. Die Genauigkeit der bestimmten Phasen kann hierdurch noch wesentlich erhöht werden. Eine wichtige Vorraussetzung für das Gelingen dieser Methode ist die Isomorphie der Kristallpackungen des nativen Proteins zu der des Schweratomderivats.

Bei einem SAD-(*single anomalous dispersion*) Experiment arbeitet man bei nur einer Wellenlänge, die in der Nähe des maximalen *f''* liegen sollte. Als anomaler Streuer sollte hier ein schwereres Element als Selen eingesetzt werden, das deutlichere ΔF_{hkl}-Werte der Friedel-Paare erzeugt. Ein Nachteil bei SAD Experimenten ist, dass bei Proteinen, die keinen Eisen-Schwefel Cluster oder ein entsprechendes Strukturelement besitzen, keine physiologische Möglichkeit zum geordneten Einbau von Schweratomen besteht. Durch Inkubation der Kristalle in schweratomhaltigen Puffern (z.B. mit

gelösten Quecksilberderivaten, die Verbindungen mit Cysteinen eingehen) lassen sich allerdings in vielen Fällen für die Phasenbestimmung brauchbare Addukte herstellen. Ein SAD-Experiment, bei dem ähnliche Schwierigkeiten bei der Phasenbestimmung wie bei einem SIR-Experiment auftreten, lässt sich mit der Methode des isomorphen Ersatzes (SIR) zu einem SIRAS Experiment kombinieren, wodurch die Zweideutigkeit der Phasenbestimmung aufgehoben wird.

3.3.8.3 ab initio Lösung

Das Phasenproblem der Röntgenstrukturanalyse kann unter Umständen für Strukturen im Größenbereich von bis zu ca. 1000 Atomen aus atomar aufgelösten Datensätzen durch direkte Methoden *ab initio* gelöst werden (Usón & Sheldrick 1999). Die notwendige Phaseninformation wird hierbei aus der Analyse der Reflexintensitäten abgeleitet.
Die zum Zeitpunkt dieser Arbeit größte, mit Direkten Methoden gelöste Struktur ohne Schweratome ist Lysozym mit 1001 Nichtwasserstoff-Atomen. Die Grenze für die Auflösung der Daten liegt für Lösungen mit Direkten Methoden bei etwa 1.2 Å, in Ausnahmefällen bei 1.4 Å.

3.3.8.4 Molekularer Ersatz (molecular replacement, MR)

Bei der Methode des molekularen Ersatzes wird eine bereits bekannte Struktur als Suchfragment zur Phasenbestimmung eingesetzt (Rossmann 1990, Lattmann 1985). In der Praxis werden hierzu die Pattersonfunktion des Suchfragmentes und die der zu lösenden Struktur in systematisch veränderten Orientierungen überlagert. Die Pattersonfunktion stellt eine Fouriersummation mit den aus dem Beugungsexperiment gemessenen Intensitäten als Koeffizienten ohne zugehöriger Phaseninformation dar:

$$P(uvw) = \frac{2}{V} \sum_{hkl} |F(hkl)|^2 \cos[2p\,(hu + kv + lw)]$$

Hierbei bezeichnen u, v und w die relativen Koordinaten der Einheitszelle und V das Zellvolumen. Mit dem Term $|F(hkl)|^2$ gehen nur die gemessenen Reflexintensitäten in die Dichtefunktion $P(uvw)$ ein. Die *peaks* in der Pattersonfuntion sind mit interatomaren

Abständen in der Zelle korreliert. Wenn ein Atom die Zellkoordinaten (x_1, y_1, z_1) besitzt und ein zweites die Koordinaten (x_2, y_2, z_2), tritt in der Pattersonfunkion ein peak für (x_2-x_1, y_2-y_1, z_2-z_1) auf. Die Höhe dieses *peaks* ist mit dem Produkt der Höhe der Elektronendichte-Maxima beider Atome korreliert. Bei atomarer Auflösung treten theoretisch N^2-N+1 *peaks* in der Patterson-Karte auf. Der *peak* auf dem Zellursprung ist der höchste, da er die Abstände aller Atome in der Zelle zu sich selbst repräsentiert. Für kleine Strukturen wie bestimmte Minerale ist die Pattersonfunktion oft dekonvolutierbar und kann als alleinige Hilfe zur Strukturlösung herangezogen werden. In der Proteinkristallographie wird die Pattersonfunktion außer bei *MR*-Methoden noch bei der Suche nach Schweratomen oder anomalen Streuern in Makromolekülen eingesetzt. Zum Vergleich der Patterson-Funktionen zweier Makromoleküle muss zunächst zwischen intramolekularen Vektoren (engl. *self vectors*), die aus Atomabständen innerhalb des Suchfragments resultieren und intermolekularen Vektoren (engl. *cross vectors*) zwischen den einzelnen Proteinen in der Einheitszelle unterschieden werden. Die Regionen der intramolekularen Vektoren in zwei Patterson-Funktionen von sehr ähnlichen Makromolekülen sollten also fast identisch sein. In der folgenden Rotationsfunktion ist die optimale Überlagerung zweier Pattersonfunktionen bei maximalem *R(fwk)* erreicht.

$$R(fwk) = \int_U P(u) P_r(u_r) du$$

Hierbei bezeichnet u die Winkelpositionen der drei Winkel f, w und k, um die P_r relativ zu P rotiert wird. Das Volumen U wird durch die Länge der intermolekularen Patterson-Vektoren bestimmt. Bei der Rotationssuche sind vor allem die starken Reflexe von Bedeutung. Im Suchmodell werden die Seitenketten, deren Lage nicht genau vorhersagbar ist, häufig durch Serin-Seitenketten ersetzt. Die Pattersonfunktion für das Suchfragment wird in einer *P1*-Zelle mit Kantenlängen berechnet, die jeweils ca. 30 Å über dem Proteindurchmesser liegen. Hierdurch wird das Auftreten von nicht klar identifizierbaren *cross vectors* vermieden. Der Auflösungsbereich der Daten für die Rotationssuche muss experimentell für die zu lösende Struktur gewählt werden und liegt in der Regel im Bereich zwischen 3 und 10 Å. Reflexe niedrigerer Auflösung sind häufig nicht sehr gut bestimmt und können außerdem zu stark von Lösungsmittelbereichen in der Kristallpackung beeinflußt werden. Daten höherer Auflösung sind zu sensitiv für das Modell, da in ihnen schon zu viele Detailinformationen über die Struktur enthalten sind, die selbst bei sehr ähnlichen

Proteinen starke Abweichungen voneinander zeigen können. Die Rotationsfunktion findet als sogenannte Selbstrotationsfunktion auch bei der Suche nach nicht-kristallographischer Symmetrie Anwendung. Hierbei wird die Pattersonfunktion um einen bestimmten Winkel um die *k*-Achse rotiert und auf sich selbst abgebildet. Die Lage der *k*-Achse, wird dabei systematisch in *f* und *w* verlagert, für jede relative Orientierung wird die Rotationsfunktion *R(fwk)* berechnet. Wenn *k*= 180 ° beträgt, erreicht *R(fwk)* für eine bestimmte Winkelposition *f, w* genau dann einen deutlichen Maximalwert, wenn eine zweizählige kristallographische oder nicht-kristallographische Achse in der Zelle parallel zur *k*-Achse verläuft. Für die Suche nach dreizähligen Achsen wählt man entsprechend *k*= 120°, für vierzählige Achsen *k*= 90° und für sechszählige Achsen *k*= 60°. Aus der Selbstrotationsfunktion kann allerdings keine translatorische Information entnommen werden, da hier Symmetrieelemente zweiter Ordnung auf Symmetrieelemente erster Ordnung reduziert werden. In einigen Fällen ist es jedoch möglich, translatorische Informationen direkt aus der Pattersonfunktion abzuleiten.

Im Falle des *molecular replacement* folgt der Rotationssuche eine Translationssuche, bei der das Fragment in der richtigen Rotationsorientierung an die korrekte Position im realen Raum geschoben wird. Im einfachsten Fall kann hierzu eine systematische Suche durchgeführt werden, indem das Suchfragment in kleinen Schritten durch die asymmetrische Einheit translatiert wird. Zu jeder Position kann der *R*-Wert

$$R = \frac{\sum_{hkl} \left| |F(obs)| - k|F(calc)| \right|}{\sum_{hkl} |F(obs)|}$$

oder ein Korrelations-Koeffizient (CC)

$$CC = \frac{\sum_{hkl} \left(|F(obs)|^2 - \overline{|F(obs)|}^2 \right) \left(|F(calc)|^2 - \overline{|F(calc)|}^2 \right)}{\sqrt{\left[\sum_{hkl} \left(|F(obs)|^2 - \overline{|F(obs)|}^2 \right)^2 \sum_{hkl} \left(|F(calc)|^2 - \overline{|F(calc)|}^2 \right)^2 \right]}}$$

berechnet werden. Hierbei ist *k* ein Skalierungsfaktor, *F(obs)* bezeichnet die Amplitude eines gemessenen, *F(calc)* die Amplitude eines für das Modell berechneten Strukturfaktors. Der Korrelationskoeffizient hat den Vorteil, dass er unabhängig von einem Skalierungsfaktor *k* von *F(calc)* zu *F(obs)* errechnet wird.

Für die Strukturlösung mit MR wurden im Rahmen dieser Arbeit zwei Programme eingesetzt. AMoRe (Navaza 1994) führt zuerst eine systematische Rotationssuche durch, an die sich eine Translationssuche für ausgewählte Orientierungen anschließt. Das Programm EPMR (Kissinger *et al.* 1999) führt eine 6-dimensionale Suche durch und startet mit einer Population von zufälligen *MR*-Lösungen, die mittels eines evolutiven Algorithmus über eine frei wählbare Anzahl von Generationen optimiert werden. Eine korrekte MR-Lösung setzt sich von falschen durch einen deutlich niedrigeren *R*-Wert sowie einen hohen Korrelationskoeffizienten ab. Sobald die richtige Orientierung und Translation des Suchfragments in der Einheitszelle gefunden sind, wird eine Verfeinerung seiner Position als starrer Körper (engl. *rigid body*) vorgenommen, in deren Verlauf der *R*-Wert deutlich absinkt.

3.3.9 Strukturverfeinerung

Alle in dieser Arbeit behandelten Strukturen wurden mit dem Programm SHELXL verfeinert. SHELXL arbeitet mit einem Algorithmus, bei dem folgende Funktion auf einen Mimimalwert optimiert wird:

$$M = \sum \left[w\left(F(obs)^2 - F(calc)^2\right)^2 \right] + \sum \left[\frac{1}{s^2} (yt - y)^2 \right]$$

Dieses *least-squares* Verfahren ist vor allem bei sehr gut aufgelösten Strukturen (unter 2 Å) und bei einem vollständigen Startmodell, das keine gravierenden Fehler enthält, effektiv. Bei Strukturen in der Größenklasse von Proteinen muss in der Regel ein *conjugate gradient* Algorithmus angewendet werden, da ansonsten der Rechenaufwand für derzeit übliche Computersysteme zu hoch wäre. Aufgrund der nicht vollständig durchgeführten Matrix-Inversion können dann allerdings keine Standardabweichungen berechnet werden. Bei einem schlechten Verhältnis der Anzahl der gemessenen Reflexe zu den im Modell zu verfeinernden Parametern sollten andere Programme eingesetzt

werden, die zusätzliche *restraints* wie z. B Vorzugsorientierungen von Seitenketten in die Verfeinerung einbeziehen. Für Strukturverfeinerungen gegen Daten mittlerer Auflösung, für die kein gutes Startmodell zur Verfügung steht, findet das Programm REFMAC (Murshudov *et al*. 1997) Anwendung, das die *maximum likelihood* Zielfunktion implementiert hat. Bei Daten, die schlechter als 3 Å aufgelöst sind, kann das Programm CNS (Brünger *et al.* 1998) verwendet werden, dass mit *simulated annealing*- und Energieminimierungs-Algorithmen arbeitet. Während der Verfeinerung wird das Strukturmodell der Elektronendichte immer weiter angepasst. Grundsätzlich, vor allem aber bei Strukturlösungen, die mit *molecular replacement* gefunden wurden, besteht das Problem, die Elektronendichte zu überinterpretieren. Diese Gefahr steigt mit abnehmender Qualität und Auflösung der gemessenen Daten. Um diesen sogenannten *model bias* so weit wie möglich zu vermeiden, kann das globale Gütekriterium des freien Residualwertes R_{free} berücksichtigt werden (Brünger 1993). Für die Berechnung des R_{free} werden ca. 5 % der Daten eingesetzt. Wichtig hierbei ist, dass das Modell ausschließlich und über die gesamte Verfeinerung hinweg gegen die übrigen 95 % der Reflexe gerechnet wird. Im abschließenden Zyklus einer jeden Verfeinerung wird der R_{free} für die Strukturfaktoren des Modells gegen die speziell markierten 5 % Reflexe berechnet. Wenn in Folge einer Fehlinterpretation der Elektronendichte der *R*-Wert abnimmt, weil das Modell in irgendeiner systematisch fehlerhaften Form an die Fourier-transformierte Dichte von 95 % der Reflexe angepasst ist, wird der R_{free} auf dem gleichen Niveau bleiben oder steigen, da bei dieser Überinterpretation die verbleibenden R_{free}-Reflexe nicht mit einbezogen wurden. Die Differenz zwischen *R*-Wert und R_{free} sollte so klein wie möglich sein und in der Regel 5 % nicht übersteigen.

3.3.10 Modelldarstellung

Zur Darstellung des Strukturmodells und der Elektronendichte wurde das Programm XFIT (McRee 1999) eingesetzt. Zur Anpassung des Modells an die gemessene Elektronendichte wurden im Rahmen dieser Arbeit zwei Dichtekarten benutzt (Read 1986). Der SigmaA-Karte liegt die Berechnung der Elektronendichte unter Einbeziehung von *F(obs)* und *F(calc)* zugrunde, wobei *F(calc)* geringer gewichtet wird. Außerdem gehen die mittleren Positionsfehler der Atome sowie die Vollständigkeit des Modells mit in die Kalkulation der Dichtekarte ein. Als zweite Dichtekarte wurde eine

Differenzdichte-Karte *F(obs)-F(calc)* zur Modellierung eingesetzt, die nicht SigmaA-gewichtet war. Bei einer aussagekräftigen Konturierung von $3s$ weist diese Darstellung auf Regionen zu hoher oder zu niedriger Elektronendichte hin.

3.3.11 Generierung von 3D-Bildern.

Die Berechnung von Strukturabbildungen mit dreidimensionalen Effekten erfolgte mit Hilfe der Programme BOBSCRIPT (Bobscript 1994-1999, Esnouf 1999) und RASTER 3D (Meritt & Bacon 1997).

3.3.12 Evaluierung von Strukturen

Für die Evaluierung der Proteinstrukturen wurden die Programme PROCHECK (Laskowsi *et al.* 1993) und WHATIF (Vriend 1990) eingesetzt. Ramachandran-*plots* wurden mit PROCHECK generiert, die Orientierungen von Asparagin- und Glutaminseitenketten wurden mit WHATIF unter dem Kriterium von Wasserstoffbrücken untersucht.

4 Ergebnisse

4.1 Aufreinigung von EETI-Varianten zu kristallographischen Zwecken

Proteine müssen zu kristallographischen Zwecken in hoher Reinheit präpariert werden. Eine übliche Substanzmasse, mit der ein Kristallisations-Experiment bei unbekannten Bedingungen begonnen werden kann, sind 10 mg Protein. Die Angabe einer Masse ist hier gebräuchlicher als die Mengenangabe in Mol, da ein messbarer Kristall eine bestimmte Mindestgröße und damit auch eine entsprechende „Mindestmasse“ besitzen muss. Bei der Produktion von sehr kleinen Proteinen wie EETI-II *in vivo* ist es schwierig, die benötigte Proteinmenge mit Standardmethoden der Molekularbiologie zu präparieren. Eine Produktionsmethode, die die benötigte Menge eines solchen Mikroproteins liefert, ist die Festphasensynthese nach Merryfield (Le-Nguyen 1989). Die Ausbildung von Sekundärstrukturelementen während der Polypeptidsynthese führt bei Cystinknoten-Inhibitoren wie EETI-II allerdings oft zu sehr niedrigen Syntheseausbeuten.

Im Rahmen dieser Arbeit wurden EETI-II (wt) und Varianten als Einschluss-Körper in *E. coli* produziert (Wentzel *et al.* 1999). Das Prinzip dieser Methode liegt in der heterologen Genexpression eines Fusionsproteins, das unmittelbar nach der Translation präzipitert und so nicht weiter in den Zellstoffwechsel eingreift. Das Fusionsprotein wird über den Zeitraum der Expression in der Zelle angereichert und kann gewöhnlich in hohen Ausbeuten präpariert werden. Ein Problem dieser Methode ist die anschließende Spaltung des Fusionsproteins und die korrekte Faltung des Zielproteins. Für die Produktion andere Mikroproteine aus der Inhibitor-Cystinknoten-Familie (ICK) sind ähnliche Strategien gewählt worden (Kojima *et al.* 1996, Bolewska *et al.* 1995).

Für die EETI-II (wt) Produktion wurde ein Gen-Konstrukt erzeugt, in dem dass *eeti*-Fragment an ein modifiziertes Gen für β-Galaktosidase angehängt wurde. Dieses Gen war so verändert worden, dass die von ihm translatierten Proteine unlöslich sind und in der Zelle zu Einschlusskörpern (engl. *inclusion bodies*) aggregieren. Zwischen den Sequenzen von β-Galaktosidase und EETI-II (wt) wurde ein Kodon für die Aminosäure Methionin inseriert, dass die Spaltung des Fusionsproteins mittels Bromzyan ermöglicht. Der Methioninrest #7 in EETI-II (wt) musste daher genetisch gegen einen Isoleucinrest substituiert werden. Das Genkonstrukt wurde über die Klonierung des

PCR amplifizierten *eeti* Fragmentes mittels *Xba*I und *EcoR*I–Verdau in den Vektor pLZPWB1 hergestellt. Die Peptidsequenz des modifizierten EETI-II (wt) Proteins war:

n-Terminus-**GCPRILIRCKQDSDCLAGCVCGPNGFCGSPHHHHHH**-c-Terminus

Die Expression erfolgte in Zellen des *E.coli* Stammes W3118, die mittels Elektroporation transformiert wurden. Zur Produktion von EETI-II Varianten wurde aus einer Glycerinkultur des plasmidtragenden Stammes eine 50 ml Kultur in dYT Medium angeimpft, der zuvor das Antibiotikum Ampicillin (100 µg/ml) zugesetzt worden war. Nach der Inkubation (ü. N., 37 °C, 180 UpM) wurde mit 25 ml dieser Kultur eine 1 l Kultur in Teriffic-Broth-Medium angeimpft (3 l Schikane-Kolben). Die Induktion der Proteinproduktion mit IPTG (1 mM Endkonzentration) erfolgte sofort nach Animpfen. Nach dem Wachstum der Kultur (ü. N., 37 °C, 150 UpM) wurde zur Ermittlung der Kulturdichte die OD_{600} bestimmt, die für die unverdünnte Kultur durchschnittlich bei 20 lag.

Die Bakterienzellen wurden mittels Zentrifugation (4 °C, 4000 UpM) geerntet und in 50 mM Tris/HCl pH 7.5 resuspendiert (30 ml). Die Einschlusskörper wurden entweder mittels *french press* Verfahren oder mittels Sonifizieren dieser Resuspension aus den Zellen freigesetzt. Das Lysat wurde abzentrifugiert (SS34, 10 min, 12000 UpM, 4°C), der Überstand wurde dekantiert.

Im ersten Reinigungsschritt wurden die Einschlusskörper im SS34-Gefäß mittels Sonifizieren in IB-Waschpuffer resuspendiert und durch Zentrifugation (SS23, 10 min, 12000 UpM, 4 °C) von in Detergenz löslichen Verunreinigungen aus dem Zell-Lysat abgetrennt. Das Pellet aus der Zentrifugation wurde erneut in IB-Waschpuffer resuspendiert. Dieses alternierende Abzentrifugieren und Resuspendieren wurde solange wiederholt, bis der Überstand nach dem Zentrifugieren klar und farblos war (bis zu fünfmal). Um das Detergenz aus dem IB-Waschpuffer zu entfernen wurde dieser Schritt noch zweimal mit Wasser (bidest.) wiederholt. Das Einschlusskörper-Präzipitat hatte eine weiße Farbe, die Ausbeute betrug im Mittel ca. 5 g Nassgewicht pro Liter Bakterienkultur.

Die Einschlusskörper wurden in 70 % (v/v) Ameisensäure (10 ml/g Einschlusskörper) gelöst. Die Lösung wurde mit Argon überschichtet. Die Spaltung des Fusionsproteins in das β-Galactosidase und das EETI-II (wt) Fragment erfolgte mittels Bromzyan, das der

Lösung zugegeben wurde (150 mg/g Einschlusskörper). Nach Rühren unter Lichtausschluss für 6 Stunden in einem 250 ml Rundkolben wurde die Lösung in ein Becherglas (1 l) überführt und mit vier Volumen eiskaltem Diethylether versetzt. Die ausgefällten Proteine wurden mittels Büchner Trichter (250 ml) unter Hausvakuum abfiltriert und mit zwei Trichtervolumen eiskaltem Diethylether gewaschen. Das nahezu trockene Präzipitat wurde bei Bedarf im Schnappdeckel Glas bei –80 °C bis zur weiteren Verwendung gelagert.

Um eine sehr homogene Fraktion von korrekt gefaltetem EETI-II (wt) zu erhalten, wurde das Präzipitat aus der Bromzyanspaltung in dn-*IMAC*-Puffer gelöst (20 ml/500 mg Trockensubstanz). Unter diesen denaturierenden (8 M Harnstoff), leicht reduzierenden (0.5 mM DTT) Bedingungen ging der gespaltene Einschlusskörper vollständig in Lösung. Nach Abzentrifugieren (SS34, 10 min, 12000 UpM) wurde der Überstand über eine zuvor mit dn-*IMAC*-Puffer äquilibrierte Nickel-Sepharose Säule gegeben (10 ml Matrixvolumen, 30 ml Gesamtvolumen). Nach dem Waschen mit dn-*IMAC*-Puffer (20 ml) wurde das gleiche Volumen *IMAC*-Puffer zur Entfernung des Harnstoffs über die Säule gegeben. Die chromatographische Trennung von EETI-II (wt) wurde mittels Stufengradienten unterschiedlicher Imidazol-Konzentrationen durchgeführt. Dazu wurden jeweils zwei Säulenvolumen 100 mM Imidazol in *IMAC*-Puffer, 500 mM *IMAC*-Puffer und 1 M *IMAC*-Puffer in der Reihenfolge der Aufzählung über die Säule gegeben. EETI-II (wt) wurde in der 500 mM Imidazol Fraktion eluiert. Die Elution von EETI-II (wt) wurde mit Hilfe von PAGE-Gelen überprüft, hierzu wurden Tris/Tricine-Gele verwendet (Abb. 10):

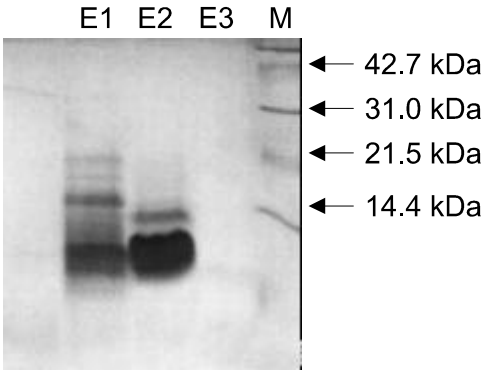


Abb. 10: Ausschnitt aus einem PAGE-Gel von EETI-II (wt) zur Kontrolle der *IMAC*-Reinigung. E1: *IMAC*-Elutionsfraktion mit 100 mM Imidazol, E2: *IMAC*-Elutionsfraktion mit 500 mM Imidazol, E3: *IMAC*-Elutionsfraktion mit 1000 mM Imidazol. M: Marker (*low range*). Rechts sind die relativen Molekulargewichte der Proteinbanden aus dem Längenstandard-Marker eingetragen.

Da eine partielle Oxidation der freien Cysteine von EETI-II (wt) während der einzelnen Aufreinigungsschritte nicht vollständig ausgeschlossen werden konnte, wurde die Elutions-Fraktion mit Harnstoff (8 M Endkonzentration) und DTT (200 mM Endkonzentration) versetzt. Nach der Inkubation (30 min, RT) wurde das reduzierte EETI-II (wt) über eine RP-Säule mittels *HPLC* gereinigt. Die Aufreinigung erfolgte über eine halb-präparative Säule (Waters µ-Bondapak C18, 3,9 x 150 mm). Es wurde ein kontinuierlicher Gradient zwischen Wasser (bidestilliert, 0.1 % TFA) als hydrophilem Laufmittel A und Acetonitril (0.1 % TFA) als hydrophobem Lösungsmittel B gefahren. Der Gradient lief von 10 % B nach 37 % B in 16 Minuten, 1 ml pro Minute. Der Durchlauf des Peptids wurde über UV-Absorption bei 217 nm detektiert, das entsprechende Volumen per Hand in Eppendorf-Reaktionsgefäßen (2 ml) aufgefangen. Nach dem Abdampfen des Lösungsmittels in der *speed vac* Zentrifuge wurde das reduzierte EETI-II (wt) gefaltet. Hierzu wurde der Deckel jedes Reaktionsgefäßes mit einer Kanüle dreifach durchstochen, daraufhin wurde das trockene Lyophilisat in 1 ml 100 mM Ammoniumcarbonat pH 8.5 gelöst. Nach Inkubation über Nacht unter Schütteln bei Raumtemperatur wurde das gefaltete Polypeptid unter gleichen Bedingungen wie das ungefaltete EETI-II (wt) über *HPLC* gereinigt (Abb. 11). Nach dem Abdampfen des Lösungsmittels in der Vakuum-Zentrifuge (*speed vac*) konnte das Peptid zu kristallographischen Zwecken eingesetzt werden.

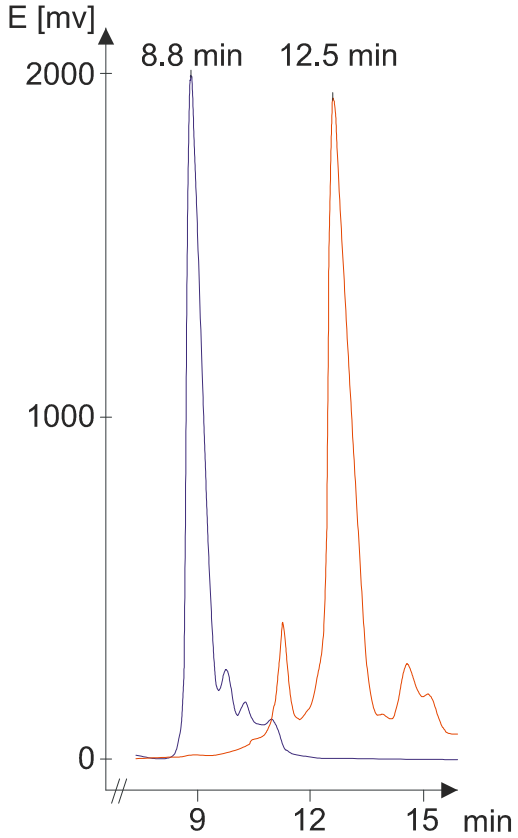


Abb. 11. Schematische Überlagerung der *HPLC*-Diagramme der chromatographischen Reinigung von reduziertem und oxidiertem EETI-II (wt). Rot eingezeichnet ist das chromatographische Profil des reduzierten Polypeptids, blau das der oxidierten Spezies. Die kleineren Nebensignale im Lauf des gefalteten Inhibitors sind vermutlich fehlerhaften Verknüpfungen der Cysteine zuzuschreiben.

4.2 Vorarbeiten zu Röntgenstrukturbestimmung von EETI-II (wt)

EETI-II (wt) wurde ausgehend von einer Proteinkonzentration von 40 mg/ml in bidestilliertem Wasser zur Kristallisation eingesetzt. Zur Suche nach Kristallisationsbedingungen wurde das *sparse matrix sampling* als *screening*-Methode angewendet (Jarmila & Kim, 1991). Als Reservoirlösungen für das *hanging drop* Verfahren (Ducruix & Griegé 1992, McPherson 1998) wurden daher 98 unterschiedliche Pufferlösungen verwendet, die in einem käuflichen *screning-kit* unter dem Gesichtspunkt eines reduzierten, aber treffsicheren Rasters zur Durchmusterung nach Kristallisationsbedingungen zusammengestellt wurden: Für die Auswahl dieser Puffer wurden zunächst nur solche Kristallisationsbedingungen aus der Literatur berücksichtigt, die bei mehreren unterschiedlichen Proteinen erfolgreich eingesetzt worden waren. Die Pufferlösungen für das *screening-kit* wurden aus dieser Menge so selektiert, dass ein möglichst breiter pH-Bereich und eine möglichst große Anzahl an unterschiedlichen Substanzen einbezogen wurden (Jarmila & Kim, 1991). Aufgrund der breiten Streuung der chemischen Zusammensetzung der Reservoirlösungen liefert das *screening* nach dieser Methode meistens einen Überblick über das Präzipitations-Verhalten des zu kristallisierenden Proteins. Ein wichtiger Vorteil des *sparse matrix sampling* ist die im Vergleich zu anderen *screening*-Verfahren geringe Menge an Protein, die für eine erste Durchmusterung auf Kristallisationsbedingungen verbraucht wird. Diese Methode ist zur Suche nach Kristallzuchtbedingungen dann besonders gut geeignet, wenn die Menge des zur Verfügung stehenden Proteins begrenzt ist.

In seltenen Fällen wachsen messbare Einkristalle direkt aus einem oder sogar mehreren *hanging drop* Ansätzen des Anfang-*screenings*. In der Regel lassen sich aber nur Hinweise auf mögliche Kristallisationsbedingungen in Form von kristallinen Niederschlägen oder ähnlichen Präzipitations-Phänomenen ableiten. Die Kristallisationsbedingungen müssen dann durch die Variation der Zusammensetzung der Reservoirlösungen in weiteren *screenings* etabliert werden.

Für die Durchmusterung der 98 unterschiedlichen Kristallisationspuffer für EETI-II (wt) wurden die „hängenden Tropfen“ aus je 1,5 µl Protein- und 1,5 µl Reservoirlösung gemischt. Die Inkubations-Temperatur betrug 4 °C. Aus dem ersten *screening* wurden

kleine Kristalle im hängenden Tropfen über der Pufferlösung der Zusammensetzung 2M Na-Formiat und 0.1 M Na-Acetat pH 4.6 erhalten. Durch Verringern der Formiat-Konzentration konnten diese Bedingungen soweit verfeinert werden, dass die Kristalle bis zu einer Größe von 0.1 x 0.2 x 0.05 mm^3 gezüchtet werden konnten (Abb. 12).

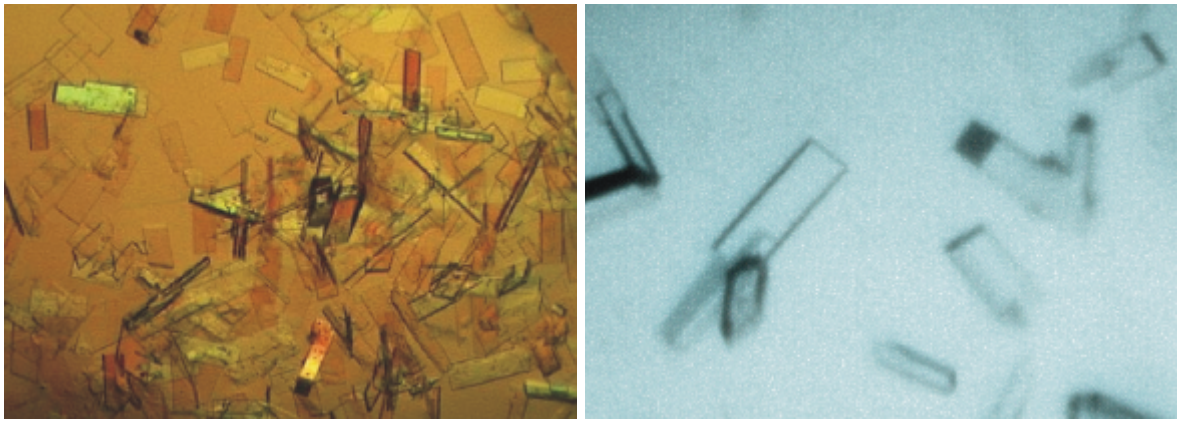


Abb. 12: Bilder von EETI-II (wt) Kristallen aus den Bedingungen: 2M Na-Formiat, 0.1 M Na-Acetat pH 4.6. Das Linke Bild ist mit polarisiertem Licht und Polarisationsfilter aufgenommen, das rechte bei normalem Durchlicht.

Die Kristalle erwiesen sich als relativ stabil gegen mechanische Einflüsse wie z. B. der Berührung mit einer Metallnadel, zeigten aber bei Erwärmung einen deutlichen Verlust an Streukraft. Da die Zugabe von Glycerol zur Mutterlauge zur Auflösung der Kristalle führte, wurde ein in der Kleinmolekül-Kristallographie übliches perfluoriertes Öl als Kryo-Protektand eingesetzt, in dem die EETI-II (wt) Kristalle bei Kühlung auf 4 °C bis zu mehrere Stunden haltbar waren. Zum möglichst schonenden Transfer der Kristalle von der Mutterlauge in den Kryo-Protektanden wurde das Deckgläschen, an dem der Tropfen mit dem Zuchtansatz hing, umgedreht auf eine bei 4 °C vorgekühlte, dicke Glaspatte gelegt. Das Öl, ebenfalls bei 4 °C temperiert, wurde vorsichtig über den Tropfen geschichtet. Der Kristall wurde mit einem *loop* aus der Mutterlauge in die Ölschicht gezogen und dort einige Sekunden leicht bewegt. Zum Zeitpunkt der ersten EETI-II (wt) Messungen war der *arc* (Abb. 9) als Montierhilfe noch nicht in der Arbeitsgruppe etabliert, die Kristalle wurden daher direkt im Kaltgasstrom gefroren. Im weiteren Verlauf dieser Arbeit konnte eine 2 molare Natrium-Malonat-Lösung als weiterer Kryo-Protektand etabliert werden. Diese Lösung hat gegenüber dem perflourierten Öl einige wichtige Vorteile: Zum einen lassen sich EETI-II (wt) Kristalle in diesem Puffer bei 4 °C praktisch konservieren. Zum anderen treten beim

Schockgefrieren der EETI-II (wt) Kristalle in Öl häufig sehr deutliche Salzringe von Formiat-Resten auf, die im Falle von Natrium-Malonat als Kryo-Protektand leichter vermieden werden können. Insgesamt wurden folgende hochaufgelöste Datensätze von EETI-II (wt) gemessen (Tab. 1):

Tab. 1: Kenngrößen der Messungen von nativen Datensätzen des Trypsin Inhibitors EETI-II (wt) an Synchrotron *beamlines*.

Messung	E.1.	E.2.	E.3.
beamline/Synchrotron	X11/DESY	BW7A/DESY	BW7B/DESY
Anzahl der Aminosäuren/Molmasse	36/3893	36/3893	36/3893
Kristallsystem	monoklin	monoklin	monoklin
Raumgruppe	*C2*	*C2*	*C2*
Zellkonstanten	a= 71.3 Å b= 82.2 Å, c= 62.7 Å β=118.7°	a= 71.3 Å b= 82.2 Å, c= 62.6 Å β= 118.7 °	a= 70.7 Å b= 81.6 Å, c= 63.0 Å β= 118.27 °
Messtemperatur	100 K	100 K	100 K
Wellenlänge	0.9057	1.379	0.8439
Unabhängige Reflexe	34775	37526	24480
Maximale Auflösung (Prozessierung)	1.7 Å	1.6 Å	1.8 Å
Vollständigkeit der Daten (%)	99.5	89.9	93.9
Durchschnittliche Redundanz	3.73	3.23	2.8
$<\|E^2-1\|>$	0.759	0.771	0.754
Rint (%)	5.48	5.16	7.18
Rs(%)	5.73	4.73	5.46
Mittleres *I/s(I)*	11.39	13.83	11.92

Für die Datensätze E.1. und E.2. wurden die EETI-II (wt) Kristalle mit Paratone-Öl als Kryo-Protektanden montiert. Die Vollständigkeit der Daten ist im Fall von Messung E.2. durch Auslassung einiger Auflösungsschalen in Folge von Salzringen im Diffraktionsbild beeinträchtigt. Die mit Malonat-Puffer als Kryo-Protektanden durchgeführte Messung E.3. zeigte im Vergleich mit den anderen Messungen das sauberste Diffraktionsbild. Die zuerst gemessenen Daten E.1. und E.2. wurden mit Hilfe der Programme XDISP, DENZO und SCALEPACK prozessiert, im Fall von Datensatz E.3. wurde die graphische Schnittstelle HKL2000 verwendet.

4.3 Ansätze zu Strukturlösung von EETI-II (wt)

Die Elementarzelle von EETI-II (wt) ist mit einem Volumen von 322012 $Å^3$ auffällig groß für den nur 36 Aminosäuren umfassenden Trypsin-Inhibitor. Die hohe mechanische Stabilität der Kristalle kann als Hinweis auf eine relativ dichte Molekülpackung interpretiert werden. Der Lösungsmittelanteil und die Anzahl von Molekülen in der Elementarzelle können auf der Basis des Matthews-Koeffizienten in oftmals guter Näherung geschätzt werden. Bei statistischen Untersuchungen stellte sich heraus, dass das Verhältnis von Zellvolumen zur Gesamtmasse an Protein V_M in der Elementarzelle für die allermeisten Proteine zwischen 1.7 und 3.5 $Å^3$/Da liegt, das Maximum der Verteilung liegt bei 2.15 $Å^3$/Da (Matthews 1968, Drenth 1994). Unter der Annahme, dass das spezifische Volumen eines Proteinmoleküls bei 0.74 cm^3/g liegt, errechnet sich der Lösungsmittelanteil V_{solv} für die Elementarzelle mit (Drenth 1994):

$$V_{Solv} = 1 - \frac{1.23}{V_M}$$

Für einen Lösungsmittelanteil von 50% (V_{solv}= 0.5) würde der Matthews-Koeffizient V_M 2.46 $Å^3$/Da betragen. Bei einer dichteren Kristallpackung wie im Falle von EETI-II (wt) wäre ein Lösungsmittelanteil von ca. 30% mit einem Matthews-Koeffizienten von V_M = 1.76 $Å^3$/Da nicht ungewöhnlich. Das relative Molekulargewicht von EETI-II (wt) beträgt 3893 Da. Da die Raumgruppe *C2* vier allgemeine Lagen besitzt, errechnet sich die Anzahl *x* von Molekülen in der asymmetrischen Einheit wie folgt:

$$x = \frac{322012Å^3}{4 * 3893Da * V_M}$$

Für einen Lösungsmittelanteil von etwa 50% wären dann acht Moleküle in der asymmetrischen Einheit vorhanden, für einen Lösungsmittelanteil von ca. 30% sogar zwölf EETI-II (wt) Monomere, wobei letzteres wahrscheinlicher ist.

Auf der Basis der in E.1. gemessenen Daten waren Lösungsansätze mit Direkten Methoden aussichtslos, da eine Auflösung von 1.7 Å weit von der hierzu benötigten atomaren Auflösung von unter 1.3 Å entfernt ist. Ausgedehnte Versuche zur Lösung der

Struktur wurden mit *molecular replacement* Methoden unternommen. Ein grundsätzliches Problem bei der Methode des molekularen Ersatzes ist die Unterscheidung zwischen inter- und intramolekularen Pattersonvektoren für die Rotationssuche. Bei Makromolekülen in der Größenklasse von EETI-II (wt) ist dies besonders schwierig, da die Länge der intramolekularen Vektoren aufgrund der engen Kristallpackung in etwa der der intermolekularen Vektoren entspricht. Die Suche von zwölf unabhängigen Molekülen in der asymmetrischen Einheit ist auch bei einem sehr guten Suchmodell relativ aussichtslos, da mit einem idealen Suchfragment bei optimaler Überlagerung nur 1/12 der Streumasse richtig interpretiert wird. Der Korrelationskoeffizient kann sich damit nicht deutlich genug von solchen für falsche Lösungen abheben. Die Suche kann in solchen Fällen unter Umständen mit Informationen über nicht-kristallographische Symmetrie vereinfacht werden. Wenn die relative Rotationsorientierung der einzelnen Moleküle in der asymmetrischen Einheit zueinander bekannt ist, kann das Suchfragment über statistische Mittelung verbessert werden, was im Endeffekt in einem größeren Verhältnis von Signal zu Rauschen resultiert: Bei einer korrekten Überlagerung in der Rotationsfunktion setzen sich die Werte für den Korrelationskoeffizienten dann wesentlich deutlicher von solchen für falsche Lösungen ab.

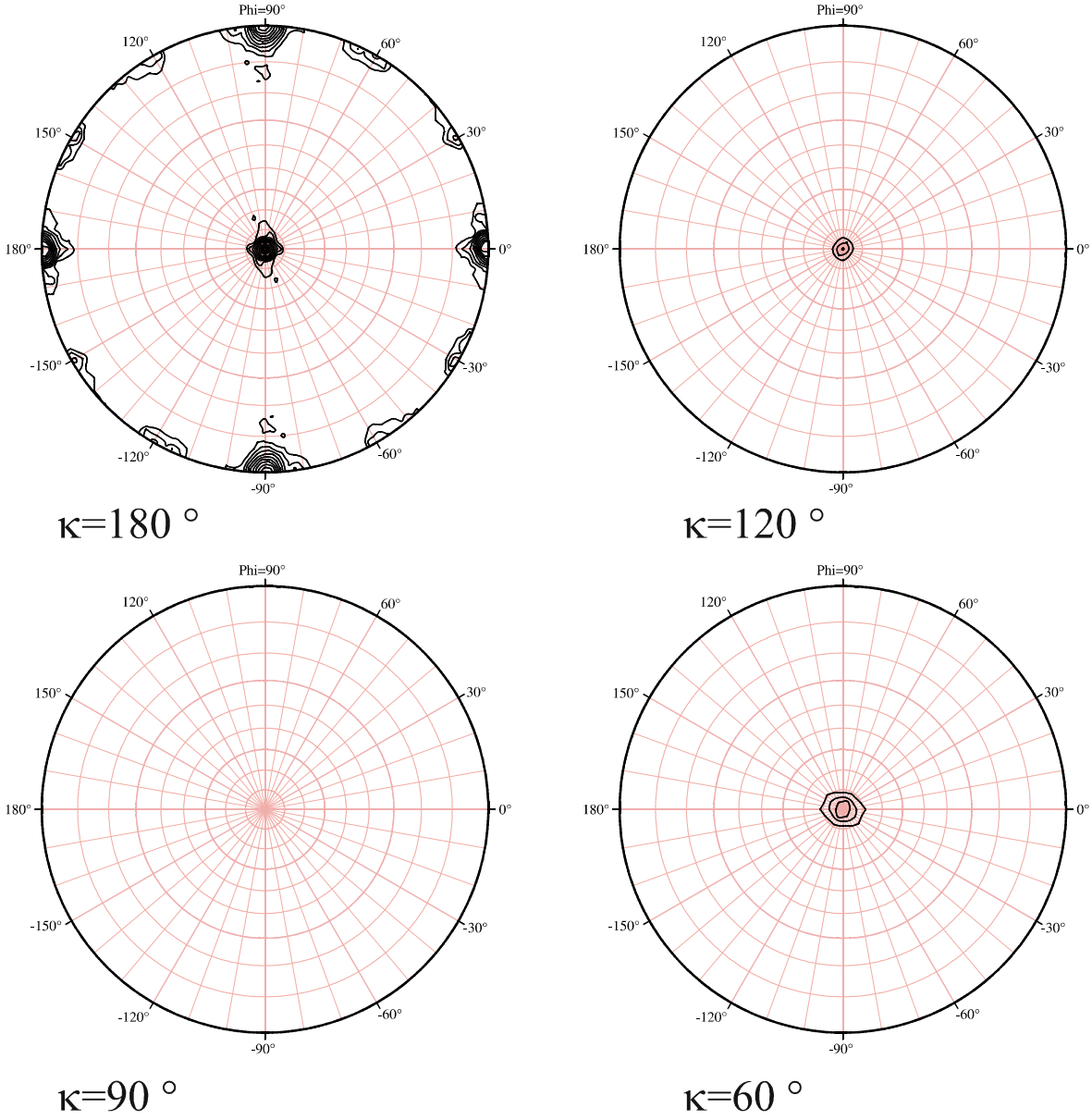


Abb. 13: Selbstrotationsfunktion von EETI-II (wt), basierend auf den Daten aus Messung E.3. Die Selbstrotationsfunktionen wurden mit folgenden Parametern berechnet: Rotationssphäre = 20.0 Å, Auflösung= 20.0-4.5 Å, Integrationsbox (g-Funktion)= 0.1. Konturierung für *k*= 180°: 500/50, für *k*= 120° und 60°: 450/50, für *k*= 90 °: 400/20. Im Mittelpunkt der Diagramme ist *w*= 0°, am Rand 90°. Die b-Achse verläuft bei *w*= 0°, die a-Achse bei *w*= 90°, *f*= 0° und c bei *w*= 90° und *f* = -118.1°.

Für die Selbstrotationsfunktion wird die Pattersonfunktion um einen bestimmten Winkel um die κ-Achse rotiert und auf sich selbst abgebildet. Die Lage der *k*-Achse wird dabei systematisch in *f* und *w* verlagert, für jede relative Orientierung wird die

Rotationsfunktion *R(fwk)* berechnet. Wenn *k*= 180° beträgt, erreicht *R(fwk)* für eine bestimmte Winkelposition *f*, *w* genau dann einen deutlichen Maximalwert, wenn eine zweizählige kristallographische oder nicht-kristallographische Achse in der Zelle parallel zur *k*-Achse liegt. Allerdings muss hierbei beachtet werden, dass in der Pattersonfunktion Symmetrieelemente zweiter Ordnung auf Symmetrieelemente erster Ordnung reduziert werden, also keine translatorische Information aus der Selbstrotationsfunktion entnommen werden kann. Für die Suche nach dreizähligen Achsen wählt man entsprechend *k*= 120°, für vierzählige Achsen *k*= 90° und für sechszählige Achsen *k*= 60°. Die unten aufgeführten Diagramme zu Selbstrotations-funktionen mit unterschiedlichen *k*-Winkeln sind Darstellungen in Polarkoordinaten. Die b-Achse der monoklinen Zelle steht senkrecht auf der Papierebene, die a-Achse verläuft in der Papierebene bei *f*= 90°.

Die Selbstrotationsfunktion bei *k*= 180° zeigt für *w*= 90° und *f* -Werte von 0°, 30°, und 60° deutliche Maxima, die auf zweizählige Achsen hindeuten. Die Signale treten, durch die zweizählige kristallographische Achse bedingt, auch bei 90°, 120°, und 150 ° auf. Für die zweizählige kristallographische Achse tritt erwartungsgemäß ein Signal bei *w*= 0° auf. Wie erwähnt, kann aus dieser Superposition der Pattersonfunktionen nicht entnommen werden, ob es sich bei den beobachteten Symmetrieelementen um reine zweizählige oder um zweizählige Achsen mit einem beliebigen Translationsanteil handelt, bzw. ob oder wie diese Achsen zueinander entlang der b-Achse verschoben sind. In der *k*= 90° Sektion tritt kein Signal auf, das auf eine vierzählige Achse hinweisen würde.

Die Selbstrotation mit *k*= 120° zeigt ein Signal für eine dreizählige Achse parallel zur Zellachse b, das bei *k*= 60° auch deutlich hervortritt und somit eine sechszählige Achse andeutet. Allerdings könnte dieser *peak* auch aus der Kombination einer dreizähligen mit der zweizähligen kristallographischen Achse resultieren. Alle im Rahmen dieser Arbeit unternommenen Versuche, die EETI-II (wt) Struktur mit Methoden des molekularen Ersatzes zu lösen, blieben trotz der Einbeziehung der Symmetrieinformation aus der Selbstrotationsfunktion erfolglos. Abb. 14 diskutiert exemplarisch drei Packungsmodelle von EETI-II (wt), die mit den Beobachtungen aus der Selbstrotationsfunktion konsistent sind:

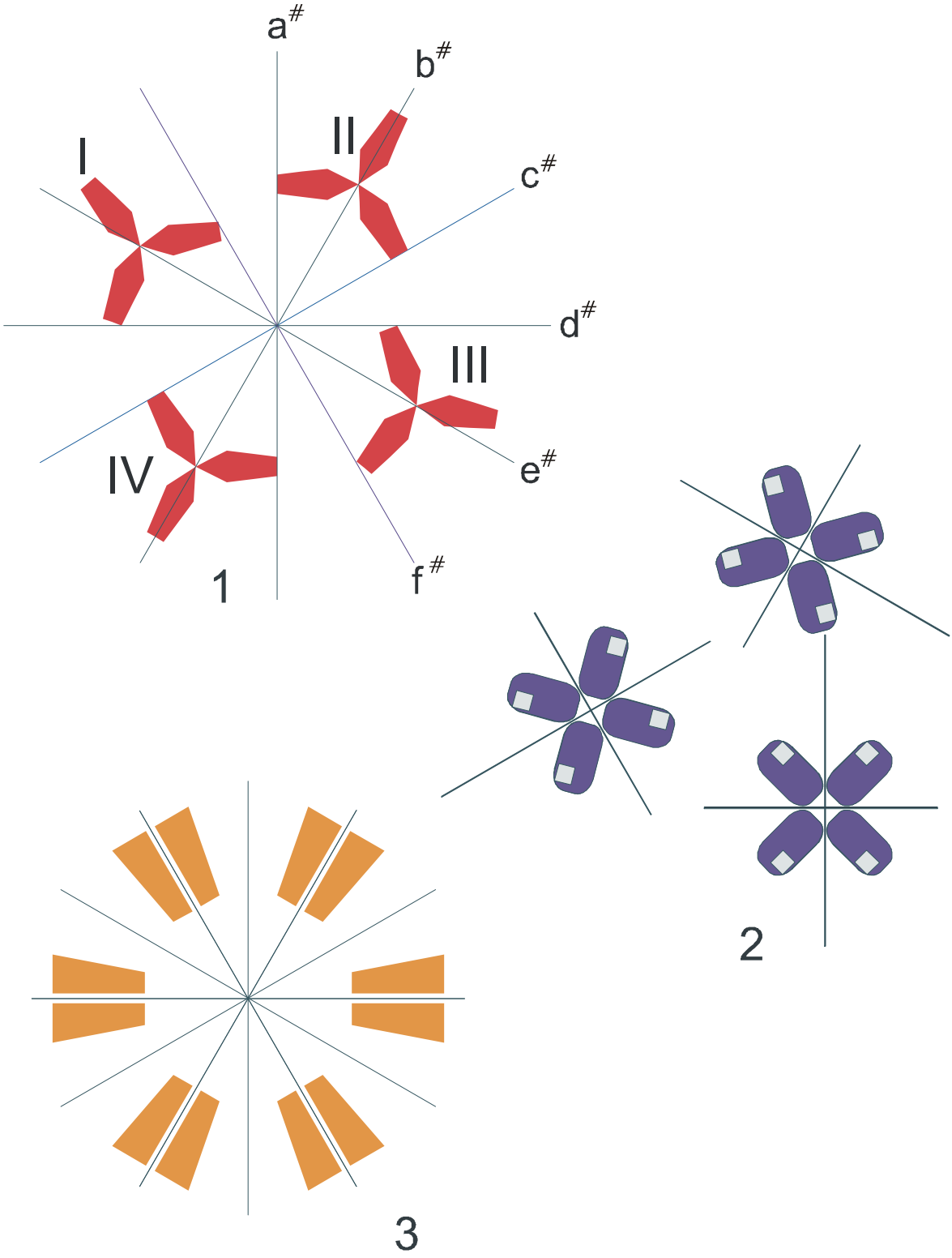


Abb. 14: Die hier aufgezeichneten Modelle für die Packung von EETI-II (wt) im Kristall repräsentieren nur drei von vielen denkbaren Möglichkeiten. Die Diskussion erfolgt im Text. Senkrecht auf der Papierebene stehende Achsen wurden nicht eingezeichnet.

4.4 Modelle für die Packung von EETI-II (wt)

Vor dem Hintergrund der Symmetrieinformationen aus der Selbstrotationsfunktion von EETI-II (wt) sollen exemplarisch drei Modelle diskutiert werden. In allen Modellen steht die b-Achse der monoklinen Zelle senkrecht auf der Papierebene.

Im Modell (1) liegen die zwölf EETI-II (wt) Moleküle in vier Trimeren vor. Durch jedes Trimer verläuft eine nicht-kristallographische, dreizählige Achse parallel zur b-Achse der Zelle. Diese dreizählige Achse wird von den Trimeren nicht perfekt erfüllt, wodurch das Signal in der k= 120° Sektion im Vergleich zu den *peaks* für die zweizähligen Achsen relativ schwach ist. Die Trimere I und III sind, lässt man translatorische Elemente aus, durch die zweizähligen Achsen $b^{\#}$, $d^{\#}$, $e^{\#}$ und $f^{\#}$ miteinander verbunden, die Trimere II und IV durch die zweizähligen Achsen $a^{\#}$, $b^{\#}$, $c^{\#}$ und $e^{\#}$. Über die aufeinander senkrecht stehenden Achsen $b^{\#}$ und $e^{\#}$ stehen nach diesem Modell alle Trimere miteinander in Verbindung, über die restlichen Achsen nur jeweils zwei Trimere. Zur Achse $a^{\#}$ wird aufgrund der auf ihr senkrecht stehenden zweizähligen kristallographischen Achse eine um 90° gedrehte Achse erzeugt, die sowohl auf der Achse $a^{\#}$ als auch auf der kristallographischen zweizähligen Achse senkrecht steht. Das selbe gilt für die Achsen $b^{\#}$ bis $f^{\#}$. Die Achse $a^{\#}$ fällt dementsprechend auf die Achse $d^{\#}$, $b^{\#}$ auf $e^{\#}$ und die Achse $c^{\#}$ auf $f^{\#}$. Dieses Modell würde erklären, warum das Signal der zweizähligen Achsen bei f= 90° stärker hervortritt als die Signale der übrigen beiden zweizähligen Achsen bei f= 30 ° und f= 60 °. Die sechszählige Achse ergibt sich in diesem Modell aus der Elimination der Translation zwischen den Trimeren II und IV.
Modell (2) erklärt die Selbstrotationsfunktion mit drei Tetrameren, die nicht auf einer vierzähligen Achse liegen, dafür aber auf zwei senkrecht aufeinander stehenden zweizähligen Achsen. Die Tetramere sind über eine dreizählige Achse miteinander verbunden, woraus sechs zweizählige Achsen im Winkelabstand von je 30° zueinander resultieren. Modell (2) erklärt nicht direkt, warum die Signale für die zweizähligen Achsen auf f = 0° und f= 90° stärker sind als für die der übrigen. Eine mögliche Begründung könnte sein, dass die zweizählige Geometrie für ein Tetramer besser erfüllt ist als für die anderen beiden. Die daraus resultierende Abschwächung des Signals für die dreizählige Achse steht in Konsens mit der Beobachtung aus der Selbstrotationsfunktion bei k= 120°.

Modell (3) geht von einem Dodekamer aus, das von je einem EETI-II (wt)-Dimer über eine sechszählige nicht-kristallographische Symmetrieachse erzeugt wird. Modell (3) ist das einzige Modell, das von der Existenz einer sechszähligen Achse ausgeht. Es erklärt nicht, warum die Signale für die zweizähligen Achsen unterschiedlich stark ausfallen, darüber hinaus ist diese Packung von Molekülen relativ unwahrscheinlich, da sie eine relativ kurze Zellkante b und lange Zellkanten a und c erwarten lässt.

4.5 Versuch der Strukturlösung von EETI-II (wt) über anomale Dispersion

Als zweiter Ansatz zur Strukturlösung wurde die Bestimmung der Schwefelatom-Positionen in EETI-II (wt) über deren anomales Signal verfolgt. Bei einer Wellenlänge von 1.54 Å hat Schwefel einen anomalen Effekt von df''= 0.56, was für die Strukturlösung eine sehr gute Datenqualität erfordert. Die Methode wurde bereits mehrfach erfolgreich in anderen Fällen angewandt, die erste über die anomale Streuung von Schwefel gelöste Struktur war Crambin (Hendrickson & Teeter 1981). Lysozym konnte ebenfalls über die anomale Streuung von Schwefel gelöst werden (Dauter *et al.* 1999). Für dieses Experiment wurden Daten in einem Auflösungsbereich von 1.54-60 Å gesammelt, die Redundanz betrug 23, das mittlere $I/s(I)$ 73. Der Residualwert R_{merge} aller Daten lag bei 4.6 %. Im Auflösungsbereich von 1.53 bis 1.56 Å betrug das $I/s(I)$-Verhältnis 10, der R_{merge} lag bei 19 %. Von EETI-II (wt) konnten auf der hauseigenen Drehanode nur Daten bis ca. 3 Å Auflösung gesammelt werden, deren Qualität weit unterhalb von der der beschriebenen Lysozymdaten lag. Aus diesem Grund wurden mehrere Messungen an EETI-II (wt) Kristallen bei einer Wellenlänge von 1.54178 Å auf dem Hausgerät (Mar345 Flächenzähler, rotierende Cu-Anode mit Osmic-*multilayer*-Optik) vorgenommen, um die Datenqualität durch eine hohe Redundanz, die über alle Messungen gemittelt bei etwa 30 lag, zu verbessern. In Tab. 2 sind einige charakteristische Größen für drei repräsentative Messungen zusammengefasst. Bei einer Auflösung von 3 Å ist es theoretisch unmöglich, zwei Schwefelatome im Cystinbrücken-Abstand von 2.02 Å voneinander aufzulösen. In der anomalen Patterson-*map* können bei dieser Auflösung prinzipiell nur Koordinaten von „Superatomen“ bestimmt werden, in denen zwei Schwefelatome zusammenfallen.

Tab. 2: Zusammenstellung der charakteristischen Werte von EETI-II (wt) Messungen zur Identifikation von Schwefellagen über anomales Signal.

Messung	A.1.	A.2.	A.3.
Röntgenquelle/Zähler	Rot. Anode/ Mar 345	Rot. Anode/ Mar 345	Rot. Anode/ Mar 345
Anzahl der Aminosäuren/Molmasse	36/3893	36/3893	36/3893
Kristallsystem	monoklin	monoklin	monoklin
Raumgruppe	*C*2	*C*2	*C*2
Zellkonstanten	a= 71.1 Å b= 82.1 Å, c= 62.8 Å β=118.5°	a= 70.8 Å b= 82.2 Å, c= 63.0 Å β= 118.4 °	a= 70.8 Å b= 82.2 Å, c= 63.0 Å β= 118.4 °
Messtemperatur	100 K	100 K	100 K
Wellenlänge	1.54178 Å	1.54178 Å	1.54178 Å
Unabhängige Reflexe	6202	6370	7060
Maximale Auflösung (Prozessierung)	3.0 Å	3.0 Å	2.9 Å
Vollständigkeit der Daten (%)	98.6	99.5	99.2
Durchschnittliche Redundanz	12.23	10.52	4.53
$<\|E^2-1\|>$	0.772	0.726	0.729
Rint (%)	7.08	9.82	8.37
Rs (%)	2.63	3.86	5.82
Mittleres *I/s(I)*	32.16	22.98	15.41

Die Signifikanz der gemessenen anomalen Differenzen wurde mit Hilfe des Programms XPREP überprüft, indem die Daten von zwei unterschiedlichen EETI-II (wt) Kristallen wie zwei Datensätze behandelt wurden, die wie bei einem MAD-Experiment bei zwei verschiedenen Wellenlängen gemessen wurden. Durch die Korrelation der anomalen Differenzen aus beiden Messungen miteinander konnte geschätzt werden, wie weit das von Messfehlern bereinigte anomale Signal tatsächlich reichte. Aus allen Messungen ließ sich ein schwaches anomales Signal bis ca. 3.6 Å bestimmen. Alle Versuche, die Schwefellagen mit Hilfe des Programmes SHELXD aus diesen anomalen Daten zu bestimmen, blieben erfolglos. Als weiterer Lösungsweg bot sich die Generierung eines Schweratomderivates an. Zur Herstellung eines potentiellen Jod-Derivates wurden

EETI-II (wt) Kristalle bei 4 °C für 1 Minute in einer Lösung inkubiert, die neben 1 M Kalium-Jodid 2 M Natrium-Malonat enthielt. Nach Ablauf der Inkubationszeit wurden die Kristalle umgehend in einem *loop* schockgefroren und montiert. Offensichtlich führte die Inkubation in der Jodlösung zu einem starken, orientierungsabhängigen Verlust an Streukraft. Die Elementarzelle, die aus einer Messung an einem Jodderivat auf dem Hausgerät (Mar345, rotierende Cu-Anode) mit a= 70.64 Å, b= 83.14 Å, c= 60.523 Å, β= 116.63° bestimmt wurde, zeigte eine zu starke Abweichung von allen bestimmten nativen Zellen, um eine Skalierung für ein SIR-Experiment durchführen zu können. Da Jod bei der Wellenlänge 1.50 Å einen anomalen Effekt von *df''*= 6.541 Elektronen besitzt, sollten die Jod-Lagen im Kristall auch aus einem SAD-Experiment bestimmbar sein. Zum Abschluss dieser Arbeit konnten an Jod-Derivaten nur Daten bis zu einer Auflösung von 4 Å gemessen werden, was für eine Phasenerweiterung nicht ausreicht. Weitere Datensammlungen und Experimente zur Erzeugung von Jod-Derivaten werden derzeit durchgeführt.

4.6 Strukturbestimmung des Trypsin/EETI-II (wt) Komplexes

Der *Ecballium elaterium* Trypsin Inhibitor II, EETI-II (wt), wurde wie eingangs beschrieben aufgereinigt. Die EETI-II (wt) Sequenz wurde zu praktischen Zwecken gegenüber der des publizierten Polypeptids (Swiss-Prot ID P12071) leicht verändert. Zum einen wurde das Methinonin #7 gegen einen Isoleucinrest ausgetauscht, da das Fusionsprotein mit Bromzyan in das EETI-II (wt) Fragment und das β-Galaktosidase-Fragment gespalten wurde (*siehe Materialien und Methoden*). Zum anderen wurde das Polypeptid zu Zwecken der Affinitätschromatographie c-terminal um sechs Histidine erweitert. Die Peptidsequenz des hier als EETI-II (wt) bezeichneten Inhibitors war:

n-Terminus-**GCPRILIRCKQDSDCLAGCVCGPNGFCGSPHHHHHH**-c-Terminus

Zur Komplexierung von EETI-II (wt) mit Trypsin aus dem Schweinepankreas wurde der Inhibitor in einem molaren Überschuss von 1.2 : 1 mit der Serin-Protease in 100 mM Bicine pH 8.0 bei einer totalen Endkonzentration von 10 mg/ml Protein assembliert. Nach einer kurzen Inkubationszeit und Zentrifugation konnte die Lösung direkt zur Kristallisation eingesetzt werden. Die Kristallisation erfolgte nach der

hanging drop Methode (Ducruix & Griegé 1992, McPherson 1998). Zur Suche nach Kristallisationsbedingungen wurde die *sparse matrix sampling* Methode (Jarmila & Kim, 1991) eingesetzt, wobei 98 unterschiedliche Pufferlösungen aus einem käuflichen *kit* verwendet wurden. Die Tropfengrößen betrugen sowohl für die Proteinlösung als auch für die jeweilige Pufferlösung 2 µl. Erste Kristalle konnten direkt aus dem Eingangs-*screen* gewonnen werden, die Zusammensetzung der Pufferlösung war: 0.1 M Natrium-Citrat, pH 5.6, 20 % PEG 4000 und 20 % Isopropanol. Der Kristall (Abb. 15) zeigte eine oktaedrische Form und die Abmessungen von 0.2 x 0.2 x 0.2 mm^3.

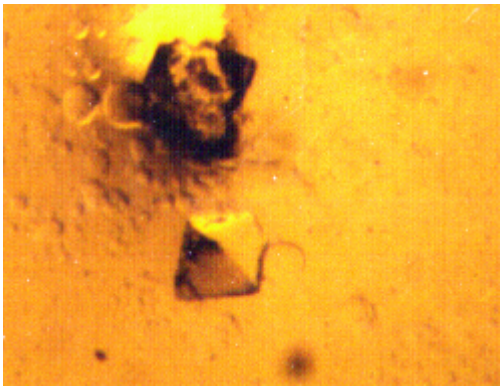


Abb. 15: Oktaedrischer Kristall von Trypsin/EETI-II (wt), gewachsen aus den Bedingungen 0.1 M Natrium-Citrat, pH 5.6, 20 % PEG 4000 und 20 % Isopropanol.

Der Kristallisationspuffer konnte durch die Zugabe von 5 % Glycerol in einen Kryo-Puffer überführt werden. Nach einer Inkubationszeit von 5 Sekunden in dieser Lösung wurde der Kristall im Stickstoffstrom des Messgerätes (Cu-Kα-rotierende Anode, *MWPC* Flächenzähler) gefroren. Bei einer Leistung von 4.5 kW wurden Daten bis zu einer Auflösung von 3.0 Å gemessen. Die Bestimmung der Elementarzelle, der Raumgruppe und der Orientierungsmatrix des Kristalls erfolgte mit dem Programm FRAMBO. Nach der Daten-Integration mit SAINT wurde die Absorptionskorrektur mit dem Programm SADABS durchgeführt, die Daten wurden mit dem Programm XPREP gemittelt. Die Strukturlösung erfolgte mit der Methode des molekularen Ersatzes (*molecular replacement*) Hierzu wurde das Programm AMoRe (Navaza 1994) verwendet, das mit einer systematischen Rotationssuche beginnt, an die sich eine Translationssuche mit einer Reihe von ausgewählten Orientierungen des Suchfragments anschließt (*siehe Materialien und Methoden*). Als Suchmodell wurde das Protease-Fragment aus der Struktur eines Komplexes von Trypsin mit *Leech-derived* Tryptase Inhibitor LDTI (PDB-ID 1LDT, Stubbs *et al.* 1997) eingesetzt. Zur Strukturlösung

wurden Daten im Auflösungsbereich zwischen 8 und 3.5 Å verwendet. Da die Raumgruppe $P4_132$ vierundzwanzig allgemeine Lagen besitzt, verlief die Rotationssuche wenig sensitiv, die Translationssuche lieferte dagegen eine sich deutlich abzeichnende Lösung. Die beste Lösung setzte sich mit einem Korrelationskoeffizienten (CC) von 50.5 und einen *R*-Wert von 45.5 deutlich von der nächst schlechteren (CC= 23.4 und *R*= 54.7) ab. Da die Qualität dieses ersten Datensatzes zur Strukturlösung, aber nicht zu einer präzisen Strukturbestimmung ausreichte, wurden weitere Kristalle für Messungen am Synchrotron gezüchtet. Eine erste Messung wurde auf der EMBL *beamline* X11 am DESY, Hamburg durchgeführt. Gegen diese Daten wurde ein vorläufiges Modell verfeinert. Der Datensatz, gegen den das endgültige Modell des Komplexes verfeinert wurde, wurde auf der *beamline* ID 14-EH2 am ESRF, Grenoble gesammelt. In Tab. 3 sind einige charakteristische Größen zu diesen Messungen aufgeführt:

Tab. 3: Kenngrößen zu Messungen des Trypsin/EETI-II (wt) Komplexes.

Messung	W/t.1.	W/t.2.	W/t.3.
Röntgenquelle/ Zähler	Rot. Anode/ *MWPC*	X11 (DESY)/ MAR 345	ID14-EH2 (ESRF)/ Mar CCD
Anzahl Aminosäuren/ Molmasse	259/27352	259/27352	259/ 27352
Kristallsystem	Kubisch	Kubisch	Kubisch
Raumgruppe	$P4_132$	$P4_132$	$P4_132$
Zellkonstanten	a= b= c= 120.40 Å	a= b= c= 121.0 Å	a= b= c= 121.61 Å*
Messtemperatur	100 K	100 K	100 K
Wellenlänge	1.54178 Å	0.9110 Å	0.9326 Å
Unabhängige Reflexe	5463	24408	49533
Maximale Auflösung (Prozessierung)	3.2 Å	1.9 Å	1.5 Å
Vollständigkeit der Daten (%)	99.9	99.5	99.7
Durchschnittliche Redundanz	37.15	19.7	8.05
$<\|E^2-1\|>$	0.787	0.803	0.767
Rint (%)	27.5	6.95	5.32
Rs (%)	10.3	5.09	2.78
Mittleres *I/s(I)*	11.80	14.69	19.24

Zuerst standen nur die auf der *beamline* X11 gemessenen Daten zur Verfügung, zur Verfeinerung wurde das Modell aus der Strukturlösung mit AMoRe eingesetzt. Nach einem ersten Verfeinerungsschritt des Modells als starrer Körper mit SHELXL (Sheldrick 1997) wurde die Elektronendichte untersucht. Teile der Elektronendichte für den Inhibitor waren deutlich sichtbar. Zur Erleichterung der Modellierung wurden Koordinaten aus einem Komplex von Trypsin (*bovine*) mit CMTI-I (PDB ID 1PPE, Bode *et al.* 1989) dem Trypsinfragment superpositioniert. Der *Cucurbita maxima* Trypsin Inhibitor CMTI-I weist den Cystinknoten auf und gehört wie EETI-II der Familie der *squash*-Inhibitoren an. Im Bereich der Protease-Bindungsregion besitzt CMTI–I dieselbe Aminosäuresequenz wie EETI-II.

Hierbei zeigte sich, dass einige Bereiche der *SigmaA*-Karte (Read 1986) und der Differenzelektronendichte im Bereich des Inhibitors durch das CMTI-I Modell sehr gut interpretiert werden konnten. Die Übereinstimmung des überlagerten CMTI-I Modells mit den Dichte-Karten war in den Regionen der Cystine und des Interaktionsbereiches mit Trypsin so eindeutig, dass die CMTI-I Koordinaten für diesen Polypeptidabschnitt in der weiteren Verfeinerung übernommen wurden. Die Modellierung wenig eindeutiger Bereiche erfolgte unter Zuhilfenahme von *omit-maps*. Hierbei wurden die kritischen Bereiche solange von der Verfeinerung ausgenommen, bis alle übrigen Abschnitte des Komplexes modelliert worden waren. In Folge des verfeinerten Modells trat die Differenzelektronendichte in den Bereichen der *omit-maps* deutlicher hervor und konnte zuverlässiger interpretiert werden. Für Seitenketten, deren Lagen nicht eindeutig zu bestimmen waren, wurden die Besetzungsfaktoren heruntergesetzt, im Falle einer nach dem darauffolgenden Verfeinerungszyklus deutlicher hervortretenden Differenzelektronendichte wurde eine diskrete Fehlordnung auf zwei Positionen modelliert. Die Besetzungsfaktoren wurden frei verfeinert, wobei sich alternative Lagen auf die Besetzung von 1 addieren mussten. Für diejenigen Aminosäuren des Modells, die bei geringer Elektronendichte keine Differenzdichte in ihrer Umgebung aufwiesen, wurde die Besetzung für die eine Lage unter Vorraussetzung einer diffusen Fehlordnung herabgesetzt, in einigen Fällen frei verfeinert. Wassermoleküle wurden in eindeutigen Fällen direkt modelliert, als Hilfsprogramm zur Modellierung von Wasser wurde daneben das Programm SHELXWAT (Sheldrick 1997) eingesetzt. Dieses Hilfsprogramm zu SHELXL wählt aus den nicht interpretierten, höchsten Elektronendichte-*peaks* eine vorgegebene Anzahl unter dem Aspekt der Koordinationsgeometrie aus und setzt Sauerstoffatome an ihre Positionen. Wenn diese

Atome im darauffolgenden Verfeinerungszyklus ein bestimmtes B-Wert-Limit nicht überschreiten, werden sie im Modell belassen, anderenfalls gelöscht. Durch diesen iterativen Suchalgorithmus kann eine entsprechende Menge von Wassermolekülen über mehrere Zyklen modelliert werden. Nachdem der in Grenoble gemessene Datensatz zur Verfügung stand, wurde die weitere Verfeinerung gegen diese Daten vorgenommen. Die wichtigsten Daten und statistischen Gütekriterien zur Verfeinerung des Trypsin/EETI-II (wt) Komplexes sind in Tab. 4 zusammengefasst:

Tab. 4: Zusammenfassung einiger charakteristischer Kenngrößen aus der Verfeinerung des Endgültigen Modells für den Trypsin/EETI-II (wt) Komplex.

Komplex	Trypsin/EETI (wt)
Aminosäuren im Modell	253
Anzahl der Proteinatome	1861
Anzahl der Lösungsmittelatome	177
Anzahl Heteroatome	1
Daten/Parameter Verhältnis	6.04
R_{work}-Wert (R_{free}-Wert) [%]	21.18 (25.37)
R-Wert (alle Daten) [%]	21.37
Anzahl der modellierten Fehlordnungen	17
Mittlere rms-Abweichungen der Bindungslängen [Å]	0.009
Mittlere rms-Abweichungen der 1,3 Abstände [Å]	0.028
Mittlerer B-Wert der Hauptkettenatome	31.08
Mittlerer B-Wert der Seitenkettenatome	34.38
Mittlerer B-Wert der Lösungsmittelatome	44.87

Die Hauptkettenverläufe von Protease und Inhibitor sind durch die Elektronendichte klar bestimmt. Im Trypsin fallen einige kurze Abschnitte des Polypeptidrückgrats in externen *loop*-Bereichen durch hohe Temperaturfaktoren auf. Das von Trypsin koordinierte Calciumatom weist einen relativ geringen B-Werte auf und wird von negativer Differenzdichte überlagert. *Bond valence* Methoden (O'Keeffe & Brese 1991) haben gezeigt, dass aufgrund der aus der Koordinationssphäre berechneten Wertigkeit des Kations nur Calcium in Frage kommen kann. Vermutlich resultiert die negative Elektronendichte aus einer Teilbesetzung der Calciumposition. Die c-terminale α-Helix von Trypsin weist ab der Position Val #235 bis zum Carboxyterminus Asn #245

teilweise hohe B-Werte in der Hauptkette und einige umgebende Restelektronendichte-*peaks* auf. Für diese Rest- und Differenzelektronendichte konnte kein befriedigendes Modell, beispielsweise in Form einer diskreten Fehlordnung der Hauptkette, gefunden werden.

Im Verlaufe der Verfeinerung zeigte sich, dass unerwartete Dichtephänomene im Bereich der Cystinreste auftraten. Die *F(obs)-F(calc)*-Karte zeigte zwischen den Sγ–Atomen der Cystinreste eine negative Differenzdichte, wohingegen in einer anderen Position zu den Cβ-Atomen der Cystine, die dem Cβ-Oγ-Abstand der Bindung im Serin entspricht, positive Differenz-*peaks* auftraten (Abb. 16).

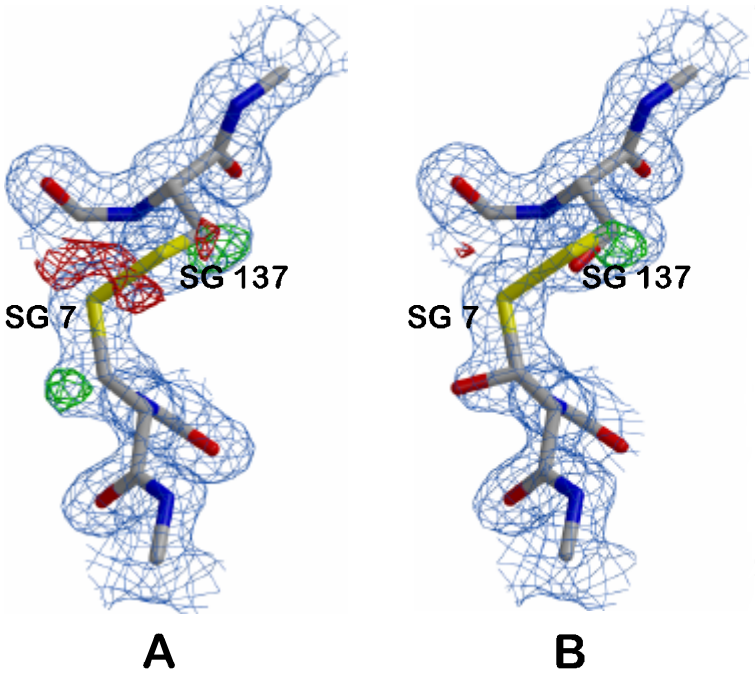


Abb. 16: (A) Als Folge von Strahlenschäden im Kristall trat negative Differenzelektronendichte auf den Schwefelbindungen zwischen einigen Cystinbücken auf. An den Cb-Atomen der Cystine traten dagegen positive Restelektronendichte-*peaks* auf. (B) Die Modellierung dieser Differenzdichte erfolgte als diskrete Fehlordnung zwischen einer Cystein- und einer Serin-Besetzung.

Nur drei der insgesamt zehn Cystinbrücken ließen sich konventionell modellieren. Im Falle von Cys #128 wurde keine Dichte für das Sγ-Atom gefunden, obwohl dieses erwartungsgemäß eine Brücke zum Sγ von Cys #232 ausbilden sollte. Die Elektronendichte der Hauptkette ist im externen *loop* #127 bis #132 relativ schlecht bestimmt, was den Schluss zulässt, dass die Cystinbrücke #128-#232 diesen Bereich der

Struktur fixiert. Die Fehlordnung der c-terminalen α-Helix von Trypsin könnte auch mit dem Bruch dieser Disulfidbindung in Zusammenhang gebracht werden.

Die beobachteten Phänomene an den Cystinpositionen erwiesen sich als konsistent mit den Beschreibungen von Schäden, die in Folge der Exposition zu Hochenergie-Synchrotronstrahlung an Proteinen auftreten können. Systematische Untersuchungen haben gezeigt, dass gerade Cystein- und Cystinreste hiervon betroffen sein können (Burmeister 2000; Weik *et al.* 2000). Die Strahlungsintensität der *beamline* ID 14-EH2 ist damit die wahrscheinlichste Ursache für die beobachtete Differenzelektronendichte an den Disulfidbrücken. Zeitgleich zum Trypsin/EETI-II (wt) Komlex wurde in der Arbeitsgruppe die Struktur von Neuraminidase aus *Salmonella typhimurium* bei einer Auflösung von 0.92 Å verfeinert (Pape *et al., to be published*). Das Auftreten desselben Differenzelektronendichte-Phänomens konnte hier als alternative Cystein-Serin Fehlordnung interpretiert werden. Die Sγ- und Cβ–Atome der betroffenen Cystinreste im Trypsin/EETI-II (wt) Komplex wurden daher als diskret fehlgeordnet auf zwei Lagen frei verfeinert, wobei die zweiten Lagen aufgrund von Bindungsabständen und Elektronendichte als Serin-Reste modelliert wurden.

4.7 Strukturbeschreibung des Trypsin/EETI-II (wt) Komplexes

Die Molekülpackung des Trypsin/EETI-II (wt) Komplexes im Kristall ist isotrop (Abb. 17), der Lösungsmittelanteil liegt bei 50 % (Matthews 1968).

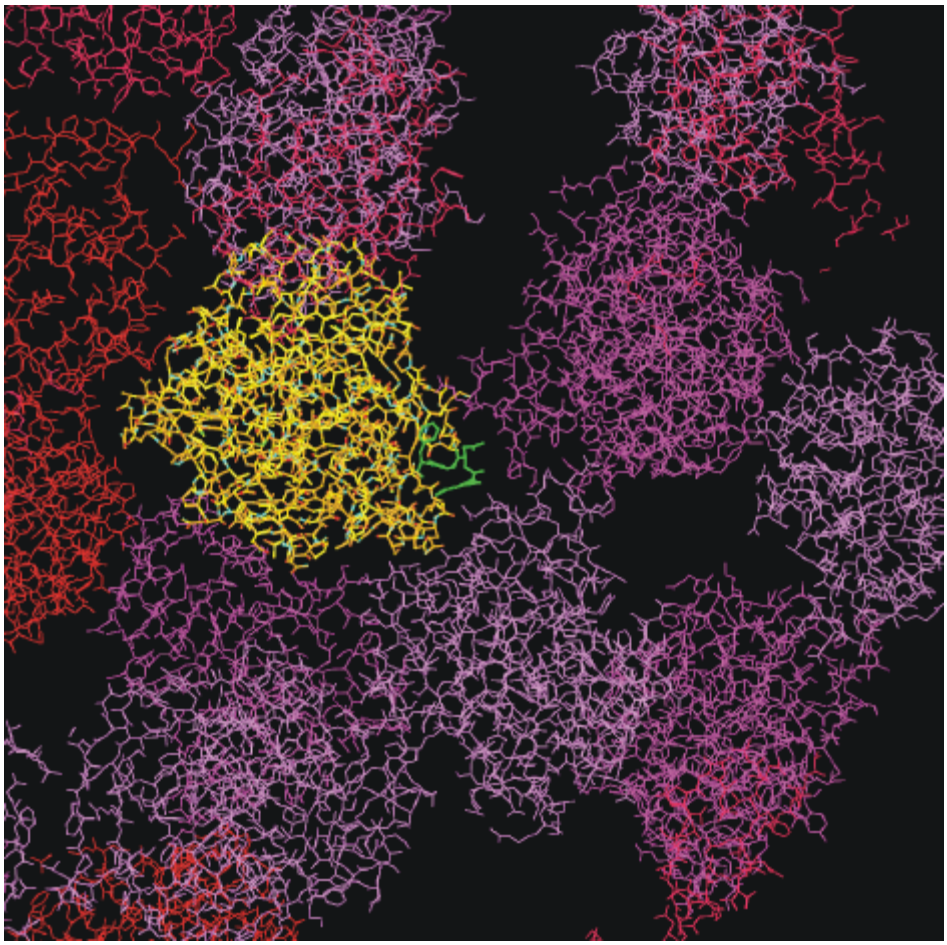


Abb. 17: Ausschnitt aus der Kristallpackung des Trypsin/EETI-II (wt) Komplexes. Die Packung von Trypsin/EETI-II (wt) füllt die Elementarzelle gleichmäßig aus. Der Sequenzbereich #22 bis #26 von EETI-II (wt) ist grün koloriert.

Alle Aminosäurereste liegen im Ramachandran-*plot* (Ramachandran & Sassiekharan 1968) in erlaubten Regionen, nur die Aminosäure Ala #244 weist eine ungünstigere Winkelgeometrie in der Hauptkette auf (Abb. 18).

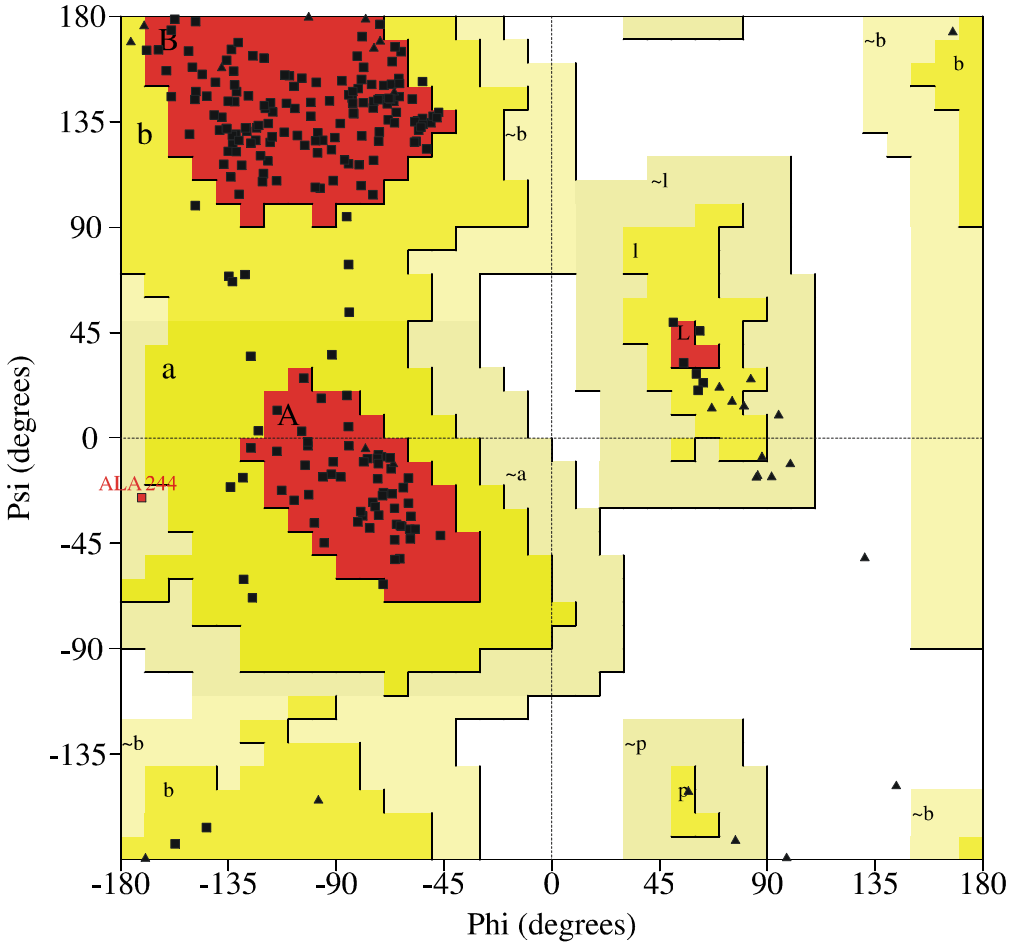


Abb. 18: Ramachandran-*plot* für die *f*/*y* Winkelverteilung im Polypeptidrückgrat des Trypsin/EETI-II (wt) Komplexes. Alle Aminosäuren liegen mit der Ausnahme von Alanin #244 (linker Rand der Grafik) in bevorzugten Regionen.

Dies könnte auf die oben beschriebene diffuse Fehlordnung der c-terminalen helikalen Region in diesem Komplex zurückzuführen sein: Der carboxy-Terminus von Trypsin (Asn #245) bildet zwei intramolekulare Wasserstoffbrücken zu den N^{ζ}-Atomen der Lysine #107 und #87 aus, wodurch die gesamte Helix auf einer definierten Position endet. Die Helix wird also aus ihrer Verdriftung auf einen Endpunkt zurückgeführt, die daraus resultierende Spannung könnte die Ursache für die ungewöhnliche Winkelgeometrie von Ala #244 sein. EETI-II (wt) weist die typischen Merkmale eines Vertreters der Inhibior-Cystinknoten-Familie auf. Abb. 19 zeigt den Verlauf der Cα-Atome von Trypsin und EETI-II (wt) im Komplex.

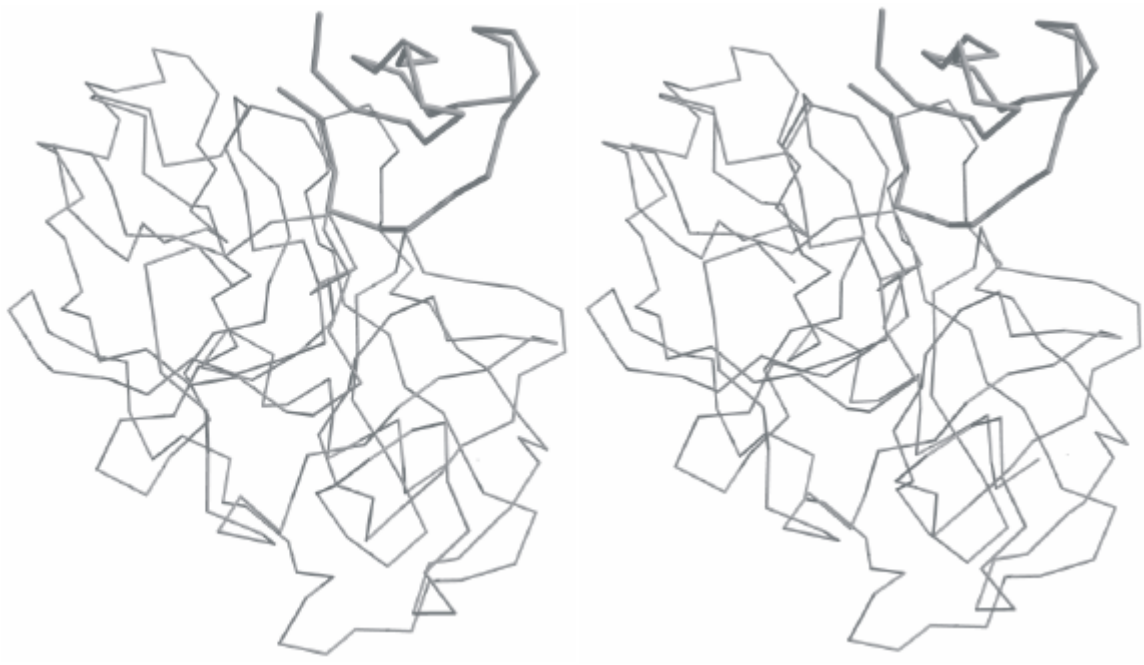


Abb. 19: Stereodarstellung des Verlauf der Cα-Atome von Trypsin (dünn gezeichnet) und dem Inhibitor EETI-II (wt) (dick gezeichnet, oben rechts) im Komplex.

Der Inhibitor sitzt über dem aktiven Zentrum der Protease, wobei im wesentlichen nur der erste *loop* von EETI-II (wt) mit dem Enzym interagiert. Die für die Trypsin/EETI-II (wt) Interaktion wichtigste Aminosäure der Inhibitorsequenz, Arg #4, ist sehr gut in der Dichte bestimmt (Abb.20 und 21). Von dieser sogenannten P1-Position (Schechter & Berger, 1967) gehen in Komplexen dieses Typs mehr als die Hälfte aller Wechselwirkungen mit der Protease aus (Bode *et al.* 1989), wodurch die Position dieses Restes sehr genau fixiert wird. Die deutlichste Interaktion ist die zum Asp #189 des Trypsins.

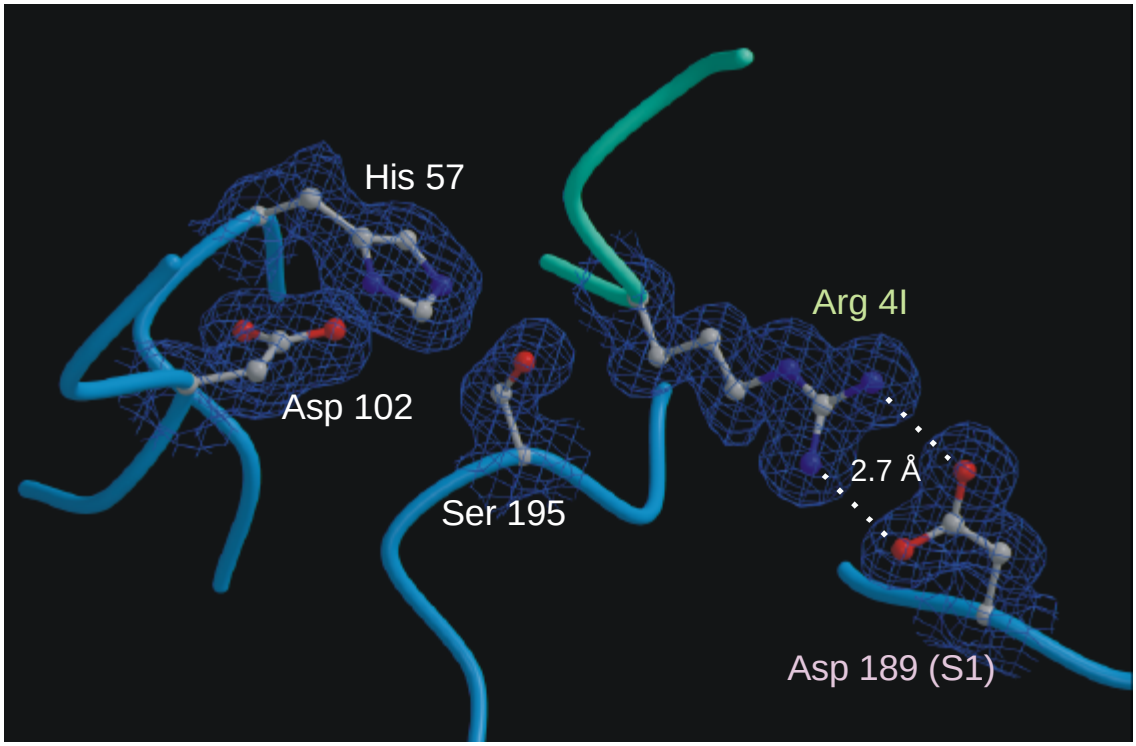


Abb. 20: Das Arg #4 von EETI-II (wt) bildet die wichtigste Interaktion zu Trypsin aus. Der Abstand zwischen den Aminogruppen des Arginins und den Carboxyl-Sauerstoffen der Seitenkette von Aspartat beträgt 2.7 Å. Durch die Bindung des Inhibitors wird die katalytische Triade von Trypsin (His #57, Asp #102 und Ser #195) blockiert.

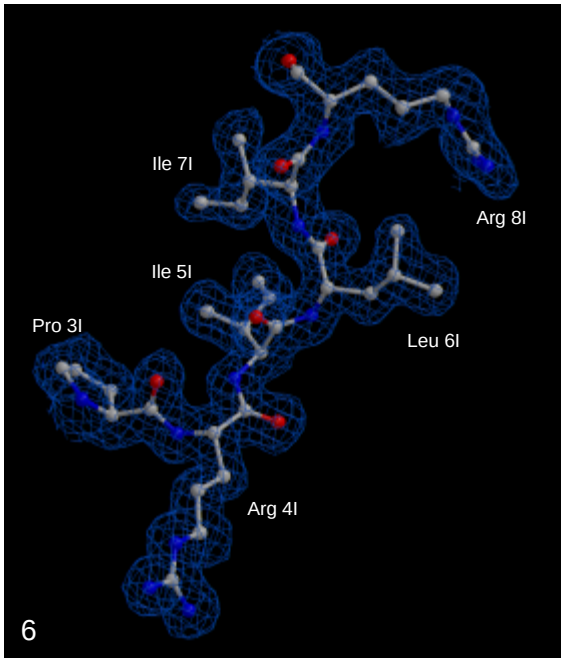


Abb. 21: Einer der in der Elektronendichte am besten definierten Bereiche der Komplexstruktur ist die Inhibitorschleife von EETI-II (wt) von Pro #3 bis Arg #8.

EETI-II (wt) zeigt das Motiv eines Cystinknotens. Die Verknüpfung der Cysteinreste ist typisch für ein Cystinknoten-Protein: Das erste Cystein ist mit dem vierten, das zweite mit dem fünften und das dritte mit dem sechsten Cysteinrest verknüpft. Daneben tritt noch ein kurzes, antiparalleles dreisträngiges β-Faltblatt und eine kurze $^{3}{}_{10}$ Helix auf (Abb. 22)

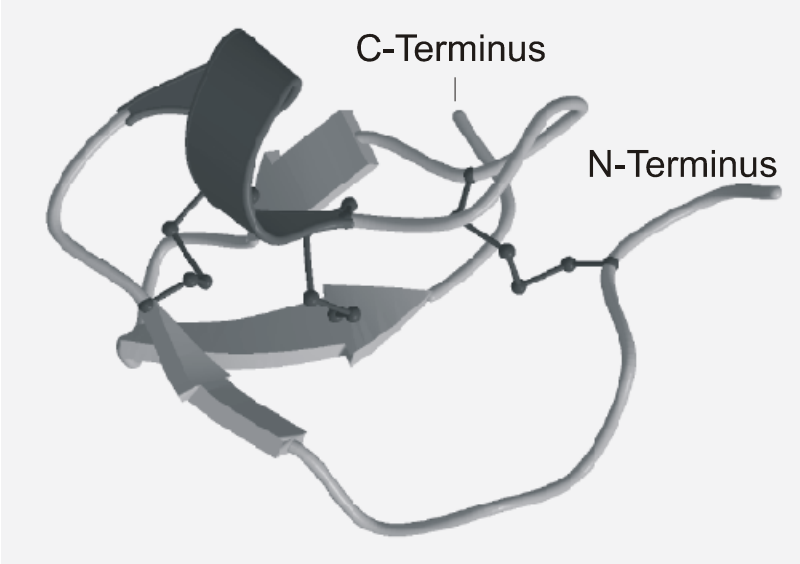


Abb. 22: Darstellung des *Ecballium elaterium* Trypsin Inhibitors als *ribbon*-Modell. Rechts im Bild liegt der n-Terminus, auf einen kurze b-Faltblatt Bereich folgt die $^{3}{}_{10}$ Helix. Bis zum c-Terminus nehmen noch zwei weitere Bereiche die Sekundärstruktur eines b-Faltblattes an, wodurch letztlich ein dreisträngiges, antiparalleles b-Faltblatt gebildet wird.

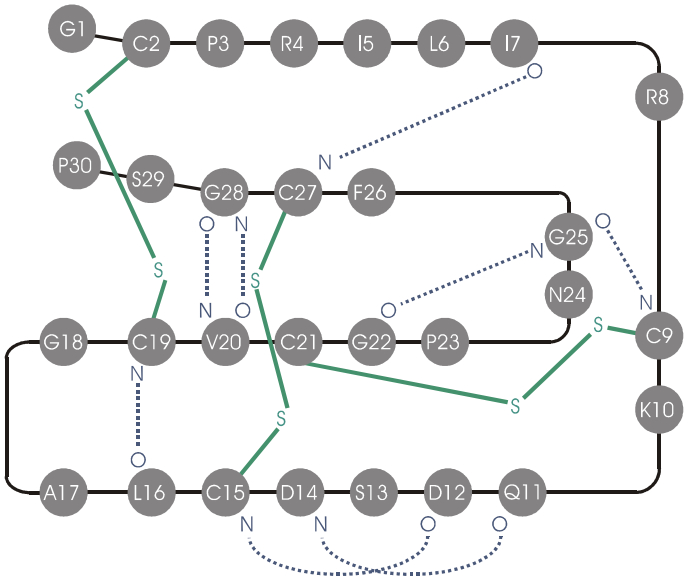


Abb. 23: EETI-II (wt) bildet in der Hauptkette eine Reihe von intramolekularen Wasserstoffbrücken aus, die zu den erwähnten Sekundärstrukturelementen führen.

Abb. 23 gibt einen Überblick über die sekundärstrukturbildenden, von Hauptkettenatomen ausgehenden Wasserstoffbrückenbindungen in EETI-II (wt). Die Aminosäure Asp #12 von EETI-II (wt) ist der einzige Aminosäurerest des Inhibitors, der zwei intramolekulare Wasserstoffbrücken von seiner Seitenkette aus knüpft. Die sechs c-terminalen Histidine sind in der Elektronendichte nicht sichtbar, in der Kristallpackung ragen sie in einen mit Lösungsmittel gefüllten Bereich. Die letzten beiden zu EETI-II (wt) gehörenden c-terminalen Aminosäuren Prolin und Serin sind dagegen gut bestimmt. Der Eingangs erwähnte GPNG-*b-turn* bildet einen Packungskontakt zu einem Symmetrieäquivalent aus, Pro #23 liegt in der Nähe der dreizähligen Achse. Der *b-turn* kann aus den Strukturkoordinaten eindeutig als Typ I *b-turn* klassifiziert werden.

4.8 Varianten von EETI-II im Komplex mit Trypsin

Von der Arbeitsgruppe PD Dr. Kolmar (Institut für Genetik und präparative Molekularbiologie, Universität Göttingen) wurden einige EETI-II Varianten aus einem in der Schleife #22-#25 erzeugten Repertoire isoliert (Wentzel *et al.* 1999), die für die Faltungsinitiation von EETI-II verantwortlich ist. Die Varianten hatten im Vergleich zum Wildtyp folgende Sequenz:

EETI-II(wt) H_2N-GCPRILIRCKQDSDCLAGCVC**GPNG**FCGSPHHHHHH-COOH
$EETI^{NEDE}$ H_2N-GCPRILIRCKQDSDCLAGCVC***NEDE***FCGSPHHHHHH-COOH
$EETI^{TNNK}$ H_2N-GCPRILIRCKQDSDCLAGCVC***TNNK***FCGSPHHHHHH-COOH
$EETI^{ERDE}$ H_2N-GCPRILIRCKQDSDCLAGCVC***ERDE***FCGSPHHHHHH-COOH

Das Selektionskriterium war die Bindung von Trypsin. Einige der isolierten EETI-II Varianten wurden von der Arbeitsgruppe Kolmar auf ihre Faltungskinetik hin untersucht, wobei deutliche Unterschiede hervortraten. Die Faltung von reduziertem EETI-II (wt) in Carbonatpuffer war nach 4 h nahezu vollständig abgelaufen, wohingegen die Variante mit der Sequenz G22N P23E N24D G25E (im folgenden als $EETI^{NEDE}$ bezeichnet) noch zu einem guten Teil als Faltungs-Intermediat vorlag, bei dem die Disulfidbrücke #2-#19 noch nicht geschlossen war. Die Variante T22G N23P K25G (im folgenden als „$EETI^{TNNK}$“ bezeichnet) lag nach vier Stunden fast vollständig

als Intermediat vor und war erst nach 16 Stunden vollständig gefaltet. Bei den Varianten $EETI^{NEDE}$ und $EETI^{TNNK}$ traten zudem falsch gefaltete Nebenprodukte auf. Die Vektoren für die Produktion der Varianten $EETI^{NEDE}$, $EETI^{TNNK}$ und $EETI^{ERDE}$ wurden von der Arbeitsgruppe Kolmar für diese Arbeit zur Verfügung gestellt. Die wichtigste Fragestellung, die mit der Untersuchung der EETI-II Varianten im Komplex geklärt werden sollte war, wie sich die stark voneinander abweichenden Faltungskinetiken in den einzelnen Röntgenstrukturen widerspiegeln.

4.9 Strukturbestimmung des Trypsin/$EETI^{TNNK}$ Komplexes.

$EETI^{TNNK}$ wurde wie EETI-II (wt) als Einschlusskörper hergestellt und analog zu EETI-II (wt) aufgereinigt. Die Peptidsequenz von $EETI^{TNNK}$ unterscheidet sich von EETI-II (wt) nur im c-terminalen *b-turn*. Die Komplexierung von $EETI^{TNNK}$ mit Trypsin wurde wie für Trypsin/EETI-II (wt) durchgeführt, allerdings wurde der Komplex nicht in 100 mM Bicine pH 8.0, sondern in 50 mM Triethanolamin (TEA) pH 7.4 bei einer totalen Endkonzentration von 10 mg/ml Protein assembliert. Die Kristallisation erfolgte in hängenden Tropfen. Die Tropfengröße betrug insgesamt 4 µl, wobei je 2 µl der Proteinlösung mit 2 µl der jeweiligen Pufferlösung vermischt wurden. Die Kristallisationsbedingungen für tetragonale Kristalle in der Raumgruppe *I4* konnten auf der Basis folgender Beobachtung etabliert werden: Wenn aus *hanging drop* Ansätzen, die mit der selben Reservoirlösung wie der für die Trypsin/EETI-II (wt)-Kristallisation angesetzt worden waren, infolge längerer Inkubation bei Raumtemperatur das Isopropanol verdampft war, bildeten sich kleine, länglich-blockförmige Kristalle in den Tropfen. Aus den Bedingungen 100 mM Natrium-Citrat, pH 5.3, 100 mM TEA pH 7.4 und 20 % PEG 4000 wurden Kristalle der Größe 0.4 x 0.1 x 0.05 mm^3 erhalten (Abb. 24).

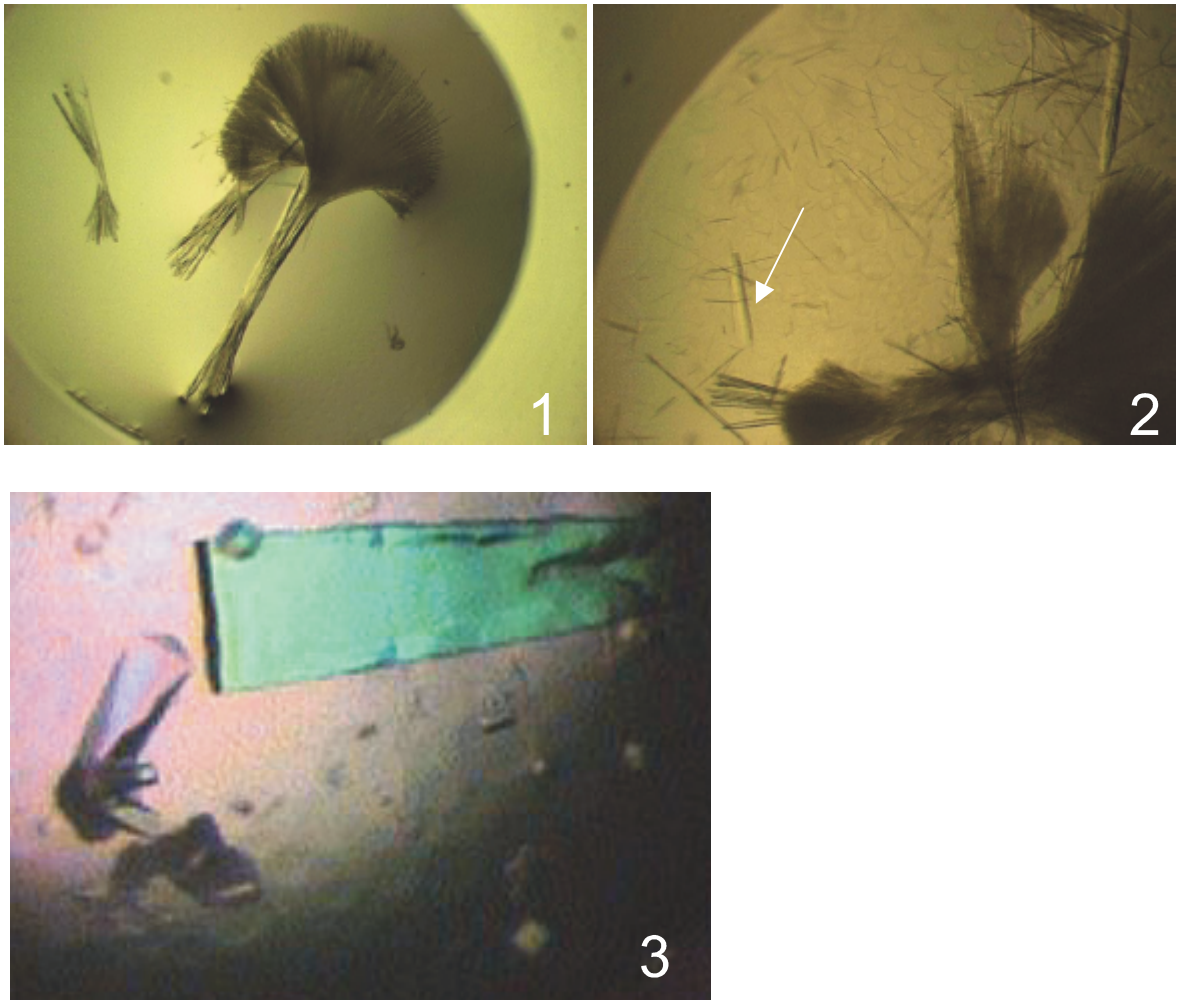


Abb. 24: Kristalle des Trypsin/EETI^{TNNK}-Komlexes. Aus den Bedingungen 100 mM Natrium-Citrat, pH 5.3, 100 mM TEA pH 7.4, 20 % PEG 4000 und 20 % Isopropanol wuchsen baumartig verzweigte Kristallgebilde (1). Nach langsamen Verdunsten des Isopropanols durch die abdichtende Fettschicht der „geschlossenen“ Kristallisationsansätze traten regelmäßige, blockförmige Kristalle (2, weißer Pfeil) auf, die bis zu einer Größe von 0.4 x 0.1 x 0.05 mm^3 heranwachsen konnten (3, rechts).

Ein geeigneter Kristall wurde in einer Glaskapillare, deren Enden mit Wachs verschlossen wurden, montiert. Für die Messung des Datensatzes wurde eine rotierende Anode (MAC Science) bei einer Leistung von 4.5 kW als Röntgenquelle und ein MAR345 *image plate* Detektor als Zähler verwendet. Die Datensammlung erfolgte bei Raumtemperatur (20 +/- 2 °C) bis 1.9 Å. Die Indizierung erfolgte mit XDISP, die Datenprozessierung mit SCALEPACK und DENZO (Otwinowski & Minor 1997). In Tab. 5 sind einige charakteristische Daten zum Kristall und zum gemessenen Datensatz zusammengefasst.

Tab. 5: Zusammenfassung einiger wichtiger Kenngrößen für die gemessenen Daten des Komplexes Trypsin/EETI^{TNNK}.

Komplex	EETI^{TNNK}
Röntgenquelle/Zähler	Rotierende Anode/MAR354
Anzahl Aminosäuren/Molmasse	259/24485
Kristallsystem	tetragonal
Raumgruppe	*I4*
Zellkonstanten	a= b= 139.86 Å c= 33.7 Å
Messtemperatur	293 K
Wellenlänge	1.54178 Å
Unabhängige Reflexe	26024
Maximale Auflösung (Prozessierung)	1.9 Å
Vollständigkeit der Daten (%)	99.4
Durchschnittliche Redundanz	6.6
$<\|E^2-1\|>$	0.737
Rint (%)	6.83
Rs	4.05
$I/s(I)$	17.48

Die Strukturlösung von Trypsin/EETI^{TNNK} erfolgte mit der Methode des molekularen Ersatzes (*molecular replacement*) unter Anwendung des Programmes EPMR (Kissinger *et al.* 1999). Das Programm EPMR startet im Gegensatz zu dem bei der Lösung des Trypsin/EETI-II (wt) Komplex eingesetzten Programmes AMoRe mit einer Population von zufälligen Orientierungen, die mittels eines evolutiven Algorithmus über eine vorgegebene Anzahl von Generationen optimiert werden. Als Suchmodell wurde das Protease-Fragment aus der Struktur des Komplexes von Trypsin mit EETI-II (wt) eingesetzt. Zur Strukturlösung wurden Daten im Auflösungsbereich zwischen 15 und 4 Å verwendet. Die Lösung besaß einen Korrelationskoeffizienten von 64.8 % und einen *R*-Wert von 36.7 %. Als weiteres Suchfragment für EPMR wurden der Inhibitor aus dem Trypsin/EETI-II (wt) Komplex eingesetzt, wobei die bereits korrekt orientierten Koordinaten der Protease festgehalten wurden. Die von EPMR vorgeschlagene Orientierung von EETI-II (wt) stimmte exakt mit der erwarteten Koordination des Inhibitors im Komplex überein. Die beiden korrekt orientierten

Komplexpartner Trypsin und EETI-II (wt) wurden im nächsten Schritt als zwei getrennte, starre Körper gleichzeitig verfeinert, wobei die Verfeinerung stabil verlief.
Da bei *MR*-Lösungen grundsätzlich das Problem des *model bias* besteht, wurden für die weitere Verfeinerung mit SHELXL einige Vorsichtsmaßnahmen getroffen. Der Bereich des TNNK-*b-turns* in $EETI^{TNNK}$ wurde solange als *omit-map* aus der Verfeinerung ausgeschlossen, bis alle übrigen Bereiche des Komplexes modelliert worden waren. Alle Abschnitte in den Polypeptidketten, die nicht eindeutig durch die Dichte bestimmt waren oder durch hohe B-Werte auffielen, wurden ebenso als *omit-maps* behandelt und schrittweise interpretiert. Im ersten Verfeinerungszyklus mit allen Parametern wurde das Modell zunächst gegen eine Auflösung von 4 Å verfeinert, die schrittweise pro darauffolgendem Zyklus um 0.1 Å bis auf 2.5 Å erhöht wurde. Hierdurch sollte die Geometrie des vorgegebenen Modells aus der *MR*-Suche am Ausgangspunkt der Verfeinerung zumindest leicht verändert werden. Die Haupketten von $EETI^{TNNK}$ und Trypsin waren in der Dichte gut bestimmt. Da die Messung bei Raumtemperatur erfolgte, fielen die B-Werte im Vergleich zum Trypsin/EETI-II (wt) Komplex durchschnittlich höher aus, vor allem in externen *loop*-Bereichen. Einige längere Seitenketten konnten daher nicht klar lokalisiert werden. Insgesamt konnten trotz der relativ niedrigen Auflösung von 1.9 Å 13 diskrete Fehlordnungen aufgrund deutlicher Indizien wie Differenzelektronendichte und hoher B-Faktoren modelliert werden. Wassermoleküle wurden wie im Falle des Trypsin/EETI-II (wt) Komplexes aufgrund sterischer Kriterien von Hand oder unter Anwendung des Programms SHELXWAT modelliert und nur dann in den nächsten Verfeinerungsschritt übernommen, wenn ihre B-Werte unterhalb von 80 $Å^2$ lagen. Der *R*-Wert der abschließenden Verfeinerung liegt mit 13.6 % viel niedriger als der des kubischen EETI-II (wt) Komplexes (21.1 %), obwohl letzterer bei höherer Auflösung (1.5 Å gegenüber 1.9 Å) gemessen wurde. Eine Ursache für diesen signifikanten Unterschied in den *R*-Werten ist vermutlich die Fehlordnung der c-terminalen α-Helix von Trypsin. Darüber hinaus besteht in der Raumgruppe $P4_132$ mit 24 allgemeinen Lagen ein höhere Wahrscheinlichkeit für Fehlordnungen der Moleküle zueinander als in der Raumgruppe *I4* mit acht allgemeinen Lagen. In Tab. 6 sind einige Daten zur Verfeinerung von Trypsin/$EETI^{TNNK}$ zusammengefasst.

Tab. 6: Zusammenfassung der wichtigsten Daten für die abschließende Verfeinerung des Modells des Trypsin/EETI^{TNNK}-Komplexes.

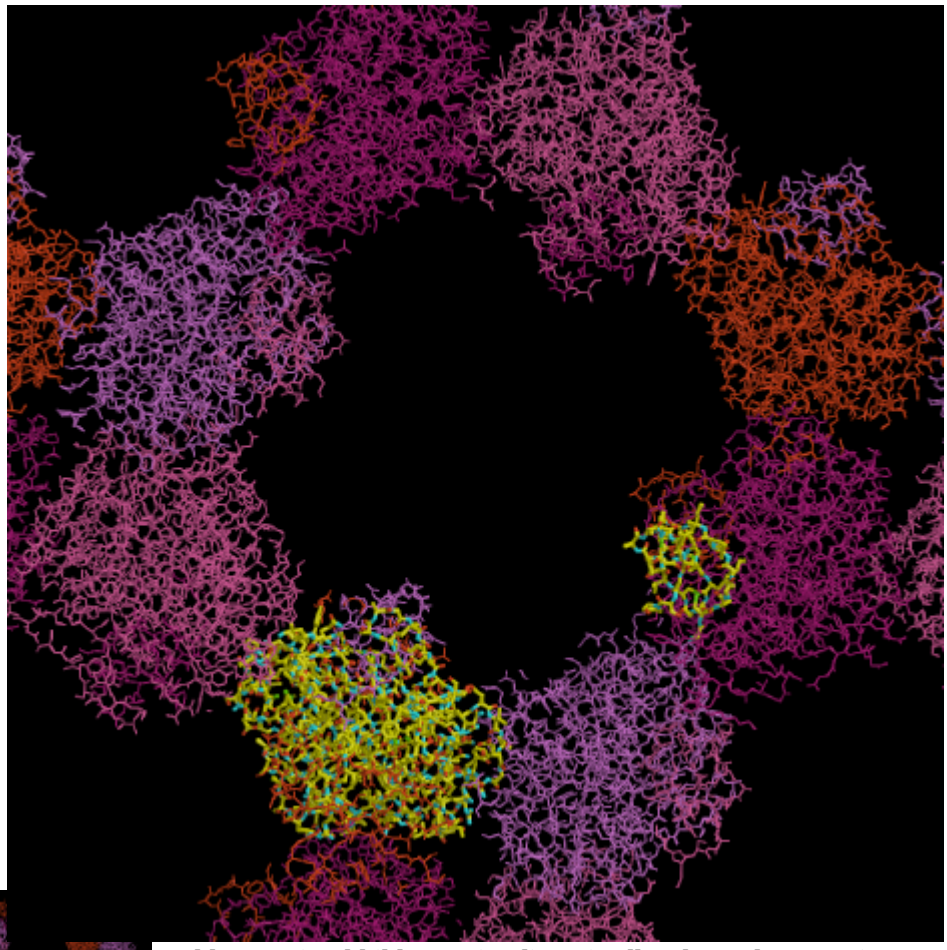
Komplex	Trypsin/EETI^{TNNK}
Aminosäuren im Modell	253
Anzahl der Proteinatome	1883
Anzahl der Lösungsmittelatome	135
Anzahl Heteroatome	1
Daten/Parameter Verhältnis	3.2
R_{work}-Wert (R_{free}-Wert)[%]	13.56 (17.07)
R-Wert (alle Daten) [%]	13.64
Anzahl der modellierten Fehlordnungen	13
Mittlere rms-Abweichungen der Bindungslängen [Å]	0.007
Mittlere rms-Abweichungen der 1,3 Abstände [Å]	0.026
Mittlerer B-Wert der Hauptkettenatome	21.48
Mittlerer B-Wert der Seitenkettenatome	24.75
Mittlerer B-Wert der Lösungsmittelatome	38.74

4.10 Strukturbeschreibung des Trypsin/EETI^{TNNK} Komplexes

Der Trypsin/EETI^{TNNK} Komplex kristallisiert in der polaren Raumgruppe *I4* mit einem Lösungsmittelanteil von 55 % (Matthews 1968). Der Kristall ist von langen, Lösungsmittel gefüllten Kanälen durchzogen, die ihn vom einen zu anderen Ende druchdringen (Abb. 25).

Die Abmessungen der Kanäle liegen bei 50 Å in der kürzesten und 80 Å in der längsten Richtung. Die Inhibitormoleküle ragen in diese Kanäle hinein, wobei der terminale TNNK-*b-turn* von EETI^{TNNK} keine Wechselwirkung mit symmetrieäquivalenten Molekülen eingeht. Alle flexibleren Hauptkettenabschnitte von EETI^{TNNK}, die in die lösungsmittelgefüllten Kanäle hineinragen, fallen durch sehr hohe B-Werte auf. Wahrscheinlich ist diese Beobachtung auf die potentiell hohe Moleküldynamik dieser *loops* zurückzuführen, die in den Röhren praktisch gelöst vorliegen. Die Beweglichkeit des Polypeptidrückgrats wird darüber hinaus noch durch die Raumtemperatur-bedingungen gefördert, unter denen die Messung durchgeführt wurde. Die Elektronendichte für die beiden terminalen EETI^{TNNK}-*b-turns* war daher für die

Seitenketten teilweise nicht vorhanden. Der Hauptkettenverlauf, der in der weiteren Diskussion dieser Komplexe eine wichtige Rolle spielt, konnte allerdings deutlich in der *omit-map* identifiziert und modelliert werden. Mit Ausnahme der unterschiedlichen Kristallpackung fallen im visuellen Vergleich der beiden Trypsinkomplexe von EETI^{TNNK} und EETI-II (wt) keine strukturellen Unterschiede im Bereich des Inhibitors auf: EETI^{TNNK} zeigt die selbe Bindungsgeometrie zu Trypsin wie EETI-II (wt), auch die intramolekularen Wasserstoffbrücken sind identisch. Im Fall von Trypsin/EETI^{TNNK} ist die c-terminale α-Helix der Protease allerdings nicht fehlgeordnet, was die Theorie der diffusen Fehlordnung in Folge von Strahlenschäden unterstützt, da das verwendete Trypsin (SIGMA) für beide Komplexe aus der selben Charge stammte.



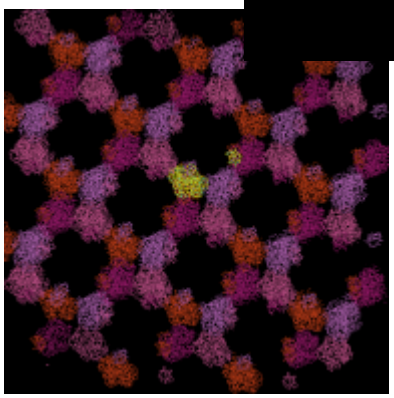


Abb. 25: Die Abbildung zeigt die Kristallpackung des Trypsin/EETI^{TNNK}-Komplexes. Großes Bild: Blick auf einen der mit Lösungsmittel gefüllten Kanäle, die den Kristall von einem Ende zum anderen durchdringen. Die Inhibitor-Moleküle ragen in diese Kanäle hinein und bilden keine Kontakte zu Symmetrieäquivalenten aus. Senkrecht zur Papierebene verläuft die vierzählige kristallographische Achse der Raumgruppe *I4*. Links: Dargestellt ist eine Übersicht der Kristallpackung.

Als weitere Erklärung hierfür kann die unterschiedliche Kristallpackung der Komplexe herangezogen werden.

Alle Aminosäuren Im Trypsin/EETI^{TNNK}-Komplex liegen in erlaubten Bereichen des Ramachandran-*plots* (Abb. 26):

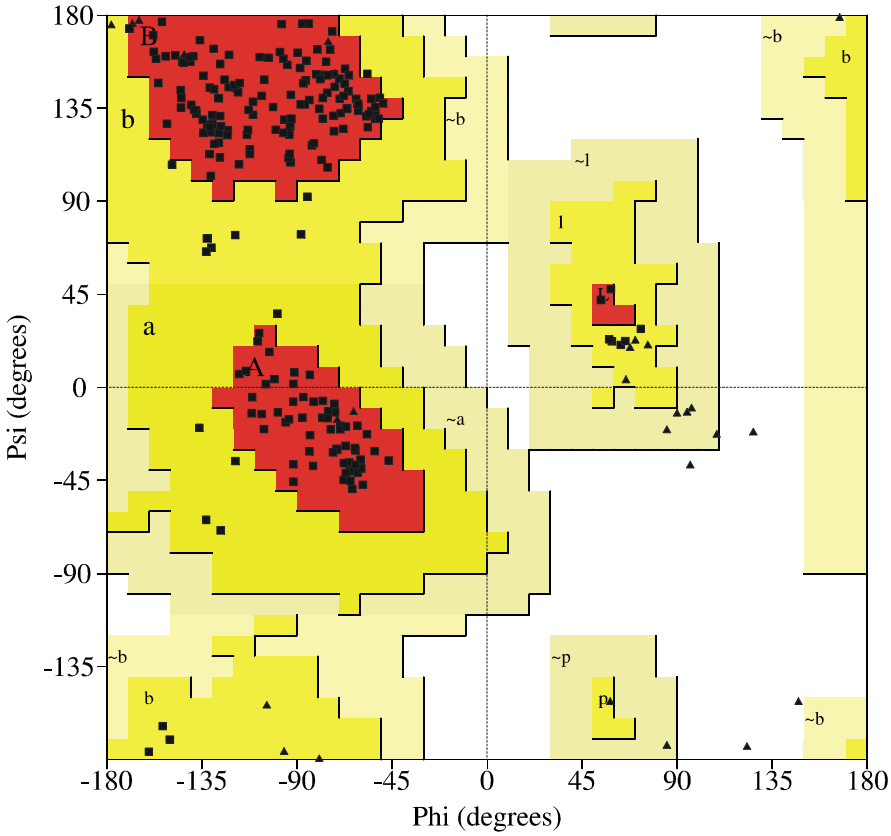


Abb. 26: Ramachandran-plot für die *f* und *y* Winkel im Verlauf des Peptidrückgrats von Trypsin/EETI^{TNNK}. Alle Aminosäurereste liegen in bevorzugten Bereichen.

4.11 Strukturbestimmung und Beschreibung der Komplexe EETI^{NEDE} und EETI^{ERDE}

Die Strukturen von zwei weiteren Trypsinkomplexen mit EETI-II-Varianten im terminalen *b-turn*, EETI^{NEDE} und EETI^{ERDE}, wurden auf der Basis von Datensätzen bestimmt, die an kleineren Kristallen der tetragonalen Form gemessen wurden. Der Kristall von EETI^{NEDE} wurde in einer Kapillare montiert, der von EETI^{ERDE} wurde für 10

Sekunden in 2 M Natrium-Malonat als Kryo-Puffer inkubiert und anschließend schockgefroren. In Tab. 7 sind einige charakteristische Daten zu diesen Messungen zusammengefasst.

Tab. 7: Zusammenstellung der wichtigsten Daten zu den Messungen der Trypsinkomplexe von EETI^{NEDE} und EETI^{ERDE}. Die Qualität der Daten ist durch sehr kleine Kristalle und, im Falle von EETI^{ERDE}, durch nicht optimierte Kryo-Bedingungen zu erklären.

Protein Komplex	porcine Trypsin/EETI^{NEDE}	porcine Trypsin/EETI^{ERDE}
Röntgenquelle/Zähler	Rotierende Anode/Mar345	Rotierende Anode/Mar345
Anzahl der Aminosäuren/ Molmasse	259/27515	259/27557
Kristallsystem	tetragonal	tetragonal
Raumgruppe	*I4*	*I4*
Zellkonstanten	a= b= 140.13 Å, c= 33.78 Å	a= b= 138.10 Å, c= 33.3 Å
Messtemperatur	293 K	100 K
Wellenlänge	1.54178 Å	1.54178 Å
Unabhängige Reflexe	6647	7233
Maximale Auflösung (Prozessierung)	3.0 Å	2.9 Å
Vollständigkeit der Daten (%)	97.6	98.9
Durchschnittliche Redundanz	5.95	4.71
$<\|E^2-1\|>$	0.728	0.730
Rint (%)	8.9	15.21
Rs (%)	12.4	18.54
Mittleres *I/s(I)*	8.11	5.27

Wie aus der Tabelle ersichtlich ist, war die Streukraft dieser Kristalle bedeutend schwächer als die der EETI^{TNNK}-Varianten, was mit ihrer wesentlich geringeren Größe (0.1 x 0.2 x 0.05 mm^3) erklärt werden kann. Daneben war der Kryo-Puffer, wie sich in weiteren Experimenten herausstellte, relativ ungeeignet für das Schockgefrieren der Komplexkristalle. Im Bereich der Hauptkette ist die *SigmaA*-Elektronendichte in beiden Strukturen relativ gut definiert, fällt in den Bereichen der mutierten *b-turns* aber nur lückenhaft aus. Beide Strukturen wurden in einem Anfangsstadium verfeinert, wobei der Bereich der Mutation durchgehend ausgeschlossen und nur als *omit-map* behandelt wurde. Es wurden keine Wassermoleküle verfeinert. Die Elementarzelle für EETI^{ERDE} ist kleiner als die der anderen *b-turn* Varianten, da hier unter Kryo-Bedingungen gemessen wurde. Bei der Überlagerung der Elektronendichtekarten von EETI^{ERDE} und

EETI^{NEDE} mit den Strukturmodellen aus EETI-II (wt) und EETI^{TNNK} fiel auf, dass der Hauptkettenverlauf der beiden gut bestimmten Modelle in die Differenz-elektronendichte der *omit-maps* von beiden neuen Varianten passte. Im Rahmen der Genauigkeit, mit der die Elektronendichtekarten bestimmt wurden, lässt sich somit aussagen, dass zwischen allen untersuchten Inhibitorkomplexen keine signifikanten Unterschiede im Bereich des Hauptkettenverlaufs des terminalen *b-turns* existieren. Vor dem Hintergrund dieser Erkenntnis wurde die genauere strukturelle Untersuchung der Komplexe von Trypsin mit EETI^{ERDE} und EETI^{NEDE} nicht weiter verfolgt.

4.12 Vergleich der Ca-Atomkoordinaten von Trypsin/EETI-II (wt) und Trypsin/EETI^{TNNK}

4.12.1 Trypsin

Die Strukturen von Trypsin sind in den beiden Komplexen Trypsin/EETI-II (wt) und Trypsin/EETI^{TNNK} nahezu identisch. Eine mit dem Programm SHELXPRO durchgeführte Überlagerung der Cα-Positionen beider Trypsinmoleküle nach dem Algorithmus der kleinsten Fehlerquadrate zeigt eine mittlere Abweichung von 0.34 Å. Auffälligere Abweichungen treten bei den Cα-Atomen von Gly #60 (Δ= 1.1 Å) und Asn #61 (Δ= 0.9 Å) auf, was auf unterschiedliche Packungskontakte zurückgeführt werden kann. Die größten Koordinatenabweichungen treten im Bereich der c-terminalen Helix auf, die im Falle des Komplexes Trypsin/EETI-II (wt) diffus fehlgeordnet ist. Asn #223 weist mit 1.26 Å die höchste Differenz der Cα-Atome auf. In Abb. 27 ist die Koordinatendifferenz der Cα-Atome der Trypsinmoleküle aus Trypsin/EETI-II (wt) und Trypsin/EETI^{TNNK} gegen die Aminosäureposition aufgetragen.

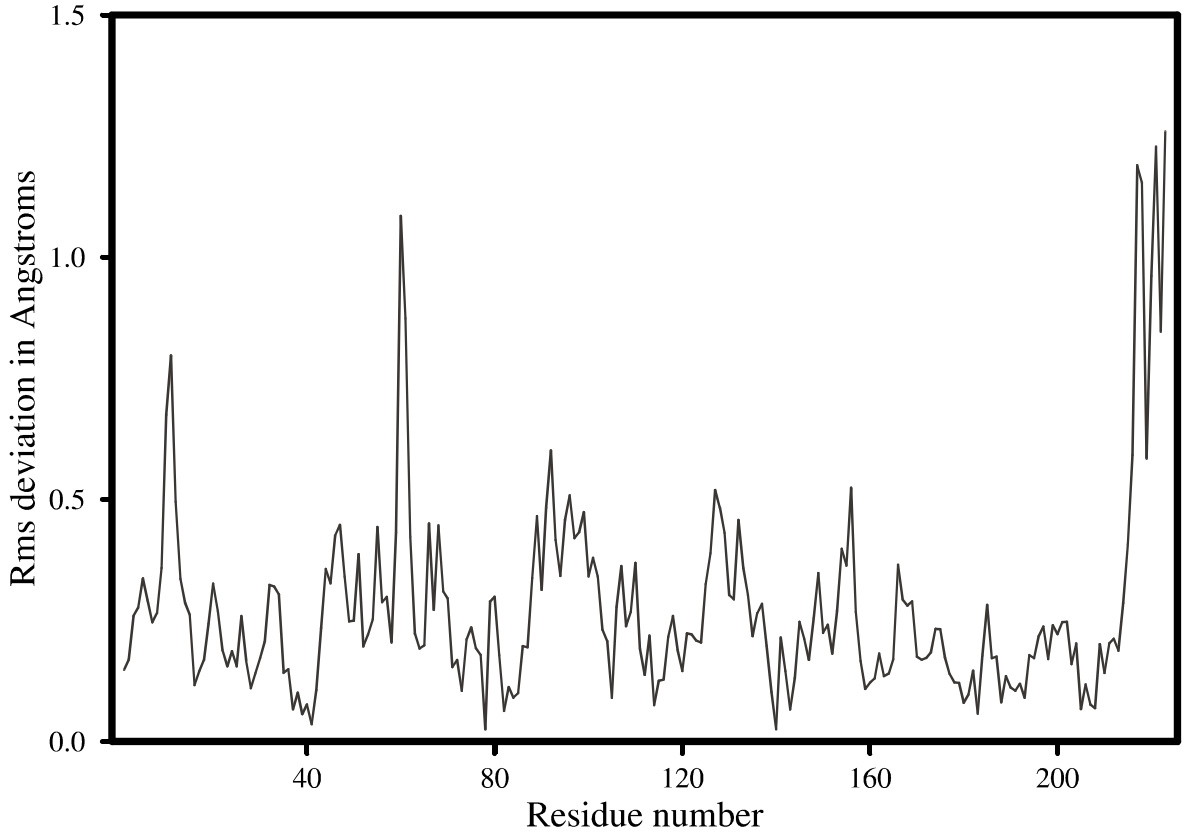


Abb. 27: Die Koordinatendifferenzen der Ca-Atome der Trypsinmoleküle aus den Komplexen von Trypsin/EETI-II (wt) und Trypsin/EETI^{TNNK} wurden gegen die Aminosäureposition aufgetragen. Alle höheren Abweichungen sind mit Ausnahme des Bereichs der c-terminalen a-Helix durch Packungskontakte zu erklären. In den hohen Positionsdifferenzen der c-terminalen Helix spiegelt sich die für den Trypsin/EETI-II (wt) Komplex beobachtete diffuse Fehlordnung wieder.

4.12.2 Vergleich der Strukturen von EETI-II (wt) und EETI^{TNNK}

Die Superposition der Cα-Atome der Strukturen von EETI-II (wt) und EETI^{TNNK} zeigt eine mittlere Abweichung von 0.26 Å. Der größte Unterschied tritt im Bereich des *b-turns* auf, der Aminsäurerest #22 weist eine Differenz von 0.58 Å zwischen beiden Strukturen auf. Diese Abweichung liegt im gleichen Größenbereich wie die, die für den c-terminalen Rest Pro #30 (0.67 Å) und für Gln #11 (0.52 Å) berechnet wurde. Gln #11 liegt in der kurzen $^{3}{}_{10}$ Helix und ist nur im Falle von EETI-II (wt) an Packungskontakten beteiligt. Die genaue Analyse von Abweichungen in der Hauptketten-Konformation an äquivalenten Positionen von EETI-II (wt) und EETI^{TNNK} zeigt, dass das Peptidrückgrat in den Bereichen der beiden *b-turns* (Abschnitte 16 bis 19 und 22 bis 25) und der

$^{3}_{10}$ Helix (Aminosäurereste 11 bis 15) der beiden Varianten sehr ähnlich verläuft (Abb. 28).

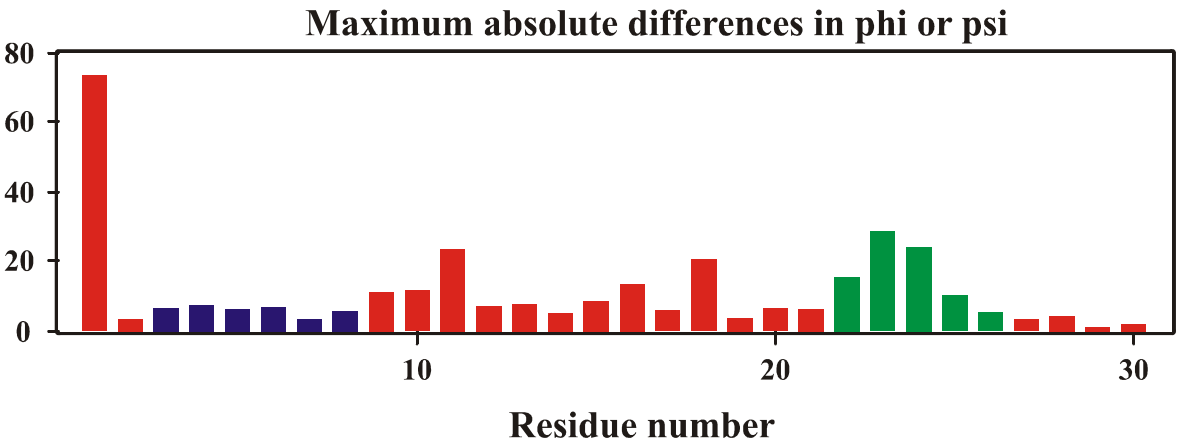


Abb. 28: Die Abweichungen zwischen den *f*- und *y*-Winkeln der beiden EETI-II Varianten wurden gegen die Position des jeweiligen Aminosäurerestes aufgetragen. Mit Ausnahme des n-terminalen Glycins #1 treten keine signifikanten Abweichungen zwischen beiden Varianten auf. Die Differenzen der *f*- und *y*-Winkel im Bereich des *b-turns* (grüne Balken) liegen in der gleichen Größenordnung wie die für die Aminosäuren #9, #10 und #11. Die Inhibitorschleife (blaue Balken) ist in beiden Komplexen sehr exakt fixiert, die Winkelabweichungen sind hier sehr gering.

Diese Bereiche sind, wie sich anhand einer B-Wert Analyse feststellen läßt, die flexibelsten Abschnitte der jeweiligen Struktur. Die Auslenkungsparameter sind hier durchschnittlich höher als in anderen Regionen der Struktur (Abb. 29).

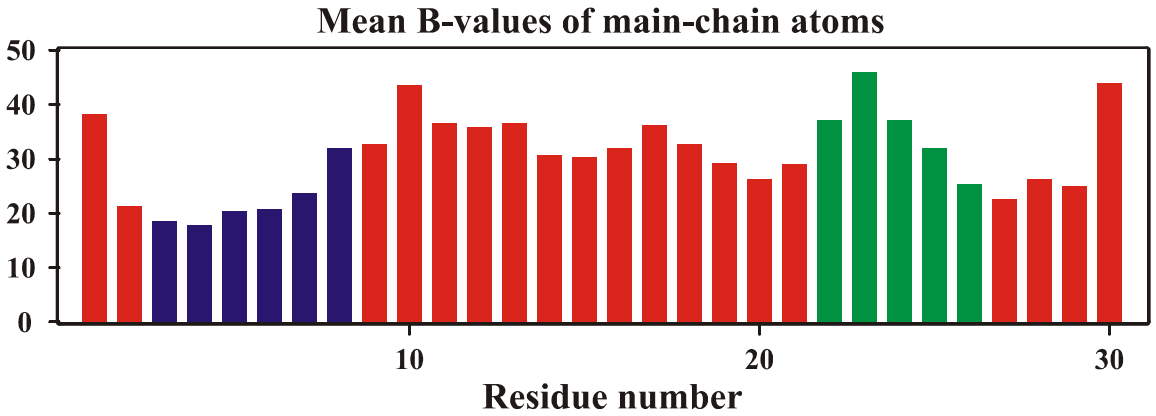


Abb. 29: Die Temperaturfaktoren der Hauptkettenatome von EETI^{TNNK} wurden gegen die Position des jeweiligen Aminosäurerestes aufgetragen. Die Bereich der beiden *b-turns* (Abschnitte 16 bis 19 und 22 bis 25) und der $^{3}_{10}$ Helix (Aminosäurereste 11 bis 15) sind, abgesehen von den Termini, die flexibelsten der Struktur.

Auf der Grundlage dieser Analyse lässt sich aussagen, dass die Koordinaten-Abweichungen im Bereich des terminalen *b-turns*, der im Falle von EETI^{TNNK} die Mutation trägt, nicht signifikant sind und daher auf die intrinsische Flexibilität der Moleküle und die unterschiedliche Packungsumgebungen zurückgeführt werden können. Abb. 30 zeigt zwei Darstellungen des terminalen *b-turns* aus der Struktur von EETI-II (wt).

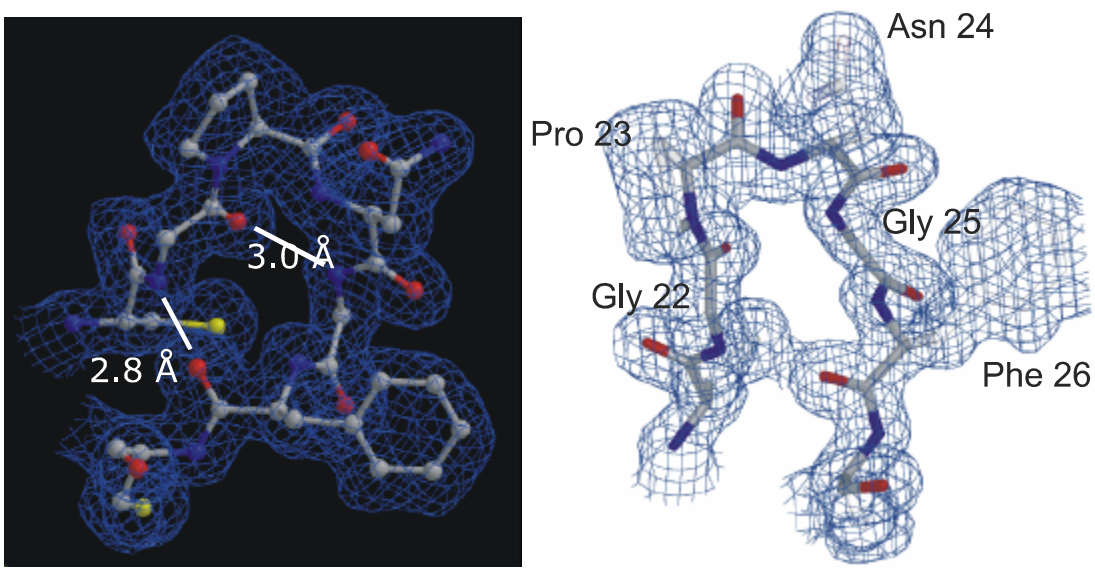


Abb. 30: Links: Der GPNG-*b-turn* von EETI-II (wt) ist ein Typ I *b-turn*. Für diese Charakterisierung werden die *f*- und *y*-Winkel der ersten vier Aminosäuren des *b-turns* herangezogen. Eine Wasserstoffbrücken-Bindung wird vom Carbonylsauerstoff des Phenylalaninrestes #26 zum Glycinrest #22 geknüpft. Rechts: Hauptkettenverlauf des Typ I *b-turns* von EETI-II (wt) in der Elektronendichte.

4.13 Vergleich von EETI-II (wt) mit den Strukturen von CMTI-I und MCTI-A

In Abb. 31 ist eine Überlagerung der *b-turns* von EETI-II (wt), EETI^{TNNK} und CMTI-I dargestellt. Die mittleren Abweichungen im Hauptkettenverlauf liegen unterhalb von 0.3 Å.

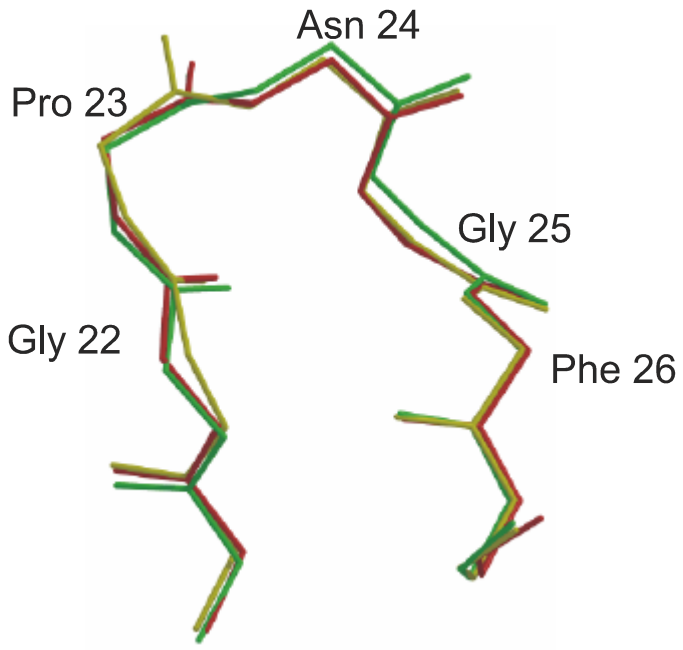


Abb. 31: Mit dem *least squares fit* Algorithmus aus dem Programm SHELXPRO wurden die Cα-Atomkoordinaten aus den *b-turns* der Inhbitoren EETI-II (wt) (gelb), EETI^{TNNK} (rot) und CMTI-I (grün) superpositioniert. Alle rms-Differenzen liegen unterhalb 0.3 Å, die Hauptkettenverläufe der drei Inhibitoren sind bei unterschiedlichen Sequenzen nahezu identisch.

Die Koordinaten von zwei anderen, vergleichbaren Trypsin-Inhibitoren, CMTI-I und MCTI-A, sind in der Protein Datenbank (PDB) hinterlegt. CMTI-I, der *Cucurbita maxima* Trypsin Inhibitor-I, wurde aus dem großen Gartenkürbis isoliert und im Komplex mit Trypsin aus dem Rind kristallisiert (Bode *et al.* 1989). MCTI-A, der *Momordica charantia* Trypsin Inhibitor-I, stammt aus den Samen der Bittermelone und wurde mit Trypsin aus dem Rinderpankreas kristallisiert (Zhu *et al*. 1999). Die Reste, die in die inhibitorische Interaktion involviert sind, können optimal superpositioniert werden (Aminosäurereste #2 bis #7 in EETI-II (wt)). Die Hauptkette dieser Inhibitoren verläuft typischerweise antiparallel zu der von Trypsin. Innerhalb der Protease-

Bindeschleife wird Cys #2 die P3- und Ile #8 die P3'-Postition zugewiesen. Die Substitution von Met #7 gegen Ile #7 hat, verglichen mit anderen Komplexen, keinen feststellbaren strukturellen Effekt auf die Trypsin/Inhibitor-Wechselwirkung. Eine Überlagerung von Cα-Atomen zwischen den komplexierten Inhibitoren EETI-II (wt), CMTI-I und MCTI-A liefert mittlere Abweichungen zwischen 0.62 Å und 0.66 Å. Abb. 32 zeigt die Überlagerung der Hauptketten für die Inhibitoren EETI-II (wt) und CMTI-I:

Abb. 32: Die Koordinaten der Ca-Atome von EETI-II (wt) (blau) und CMTI-I (rot) wurden mit Hilfe des Programms SHELXPRO superpositioniert. Die mittlere Abweichung der Koordinaten beträgt 0.62 Å.

Die Untersuchung von Differenzen zwischen äquivalenten Hauptketten-Torsionswinkeln zeigt, dass diese im Bereich der Postionen #17 und #18 am höchsten sind: Die EETI-II Varianten weisen hier im Gegensatz zu den übrigen Inhibitoren einen Pepflip auf. Die Sγ-Atome der Cytinreste können mit einer mittleren Abweichung von 0.25 Å superpositioniert werden, obwohl MCTI-A einen Rest weniger im terminalen *b-turn* besitzt (Abschnitt #22 bis #25).

4.14 Vergleich der EETI-II (wt) Röntgenstrukturen und der NMR Struktur

Zum Zeitpunkt dieser Arbeit waren in der PDB die NMR Struktur eines c-terminal um die Aminosäuren Serin #29 und Prolin #30 verkürzten EETI-II Wildtyps (Chiche *et al.* 1989) und ein Ensemble von 20 unterschiedlichen Strukturen für die Mutante EETI I5L (Nielsen *et al.* 1994) deponiert. Die 20 leicht voneinander abweichenden Lösungen der Variante EETI I5L können als Modell für die Moleküldynamik von EETI-II interpretiert werden. Die Protease-Bindeschleife zeigt die am weitesten voneinander abweichenden Konformationen zwischen allen Lösungen im Verlauf der Hauptkette. Da dieser Bereich bei der Komplexierung eindeutig fixiert ist, wurde der Polypeptidabschnitt der Aminosäuren #1 bis #7 für einen Vergleiche der Cα-Atomkoordinaten ausgespart. Die hohe Flexibilität der Inhibitorschleife ist vermutlich ein weiterer Grund, aus dem die Lösungsversuche von freiem EETI-II (wt) mit Methoden des molekularen Ersatz erfolglos blieben. Die mittlere Abweichung von 1.11 Å war viel größer als die, die aus dem Vergleich der unterschiedlichen Inhibitoren im Kristallverband berechnet wurde. Die maximalen Abweichungen in den Cα-Atompositionen im analysierten Proteinabschnitt treten im c-terminalen *b-turn* auf (Pro #23: 2.15 Å, #22 und #23 > 1.4 Å). Die Ursache hierfür kann in einer hohen Flexibilität dieser Region oder in den experimentellen Fehlergrenzen der NMR-Methode zur Bestimmung von Strukturen in Lösung liegen.

5 Diskussion

Aus dem Vergleich der Kristallstrukturen der untersuchten EETI-II Varianten untereinander und mit anderen Inhibitoren aus der *squash*-Inhibitorfamilie wird deutlich, dass der Cystinknoten und weite Teile der Hauptkette unter strukturellen Gesichtspunkten streng konserviert sind. Auch der Bereich des c-terminalen *b-turns* vom Typ I lässt sich für alle Kristallstrukturen nahezu deckungsgleich superpositionieren, obwohl einige der eingeführten Mutationen wie Prolin gegen Asparagin oder Glycin gegen Threonin theoretisch eine signifikante Änderung des Polyeptidrückgratverlaufs bedingen könnten. Die unterschiedlichen Faltungskinetiken, die für die EETI-II Varianten bestimmt wurden (Wentzel *et al.* 1999), spiegeln sich demnach nicht in der endgültigen Tertiärstruktur wieder, sondern haben ihre Ursache eher in der Zeitabhängigkeit der Ausbildung eines initialen oder intermediären Faltungskerns.

Zur weiteren Diskussion der Abhängigkeit der Tertiärstruktur von der Aminosäuresequenz der Inhibitoren wurde der Konservierungsgrad der einzelnen Aminosäurereste von EETI-II im Vergleich mit 17 anderen Inhibitoren aus der ICK-Familie qualitativ bestimmt. Es wurden nur Sequenzen aus der Swiss-Prot Datenbank (Bairoch & Apweiler 2000) verwendet, die im Bereich zwischen dem ersten und dem letzten Cystein des Polypeptidrückgrats nicht identisch waren. Die verwendeten Sequenzen waren die folgenden (Abb. 33):

```
EETI-II                          GCPRILMRCKQDSDCLAGCVCGPNGF-CG
MCEI-I                          RICPLIWMECKRDSDCLAQCICVDGH--CG
CMTI-I                          RVCPRILMECKKDSDCLAECVCLEHGY-CG
MRTI-I                          GICPRILMECKRDSDCLAQCVCKRQGY-CG
CMCTI-I, CMCT-II               QRMCPKILMKCKQDSDCLLDCVCLKEGF-CG
AND CMCTI-III
CMTI-IV AND                  HEERVCPRILMKCKKDSDCLAECVCLEHGY-CG
CMTI-III
TRYPSIN INHIBITOR A             RSCPRIWMECTRDSDCMAKCICVAGH--CG
LLDTI-I OR LLDTI-II            QRRCPRIYMECKHDSDCLADCVCLEHGI-CG
OR LLTI-III
LCTI-I                          RICPRILMECSSDSDCLAECICLEQGF-CG
BDTI-II                         RGCPRILMRCKRDSDCLAGCVCQKNGY-CG
LCTI-II                         RICPRILMECSSDSDCLAECICLEQDGFCG
CPTI-III AND CPTI-II         HEERVCPKILMECKKDSDCLAECICLEHGY-CG
CSTI-IV                         MMCPRILMKCKHDSDCLPGCVCLEHIEYCG
CVTI-I                         GRRCPRIYMECKRDADCLADCVCLQHGI-CG
MCTI-I                         ERRCPRILKQCKRDSDCPGECICMAHGF-CG
CSTI-IIB                        MVCPKILMKCKHDSDCLLDCVCLEDIGYCGVS
MCTI-II                         RICPRIWMECKRDSDCMAQCICVDGH—-CG
TGTI-I                           ICPRILMPCSSDSDCLAECICLENGF-CG
```

Abb. 33: Sequenz-*alignment* von *squash*-Inhibitoren aus der Swiss-Prot Datenbank. Die dem GPNG-*b-turn* von EETI-II entsprechenden Aminosäurepositionen sind rot koloriert. Die Swiss-Prot-Nummern und die Organismen, aus denen die einzelnen Inhibitoren isoliert wurden, sind: EETI-II P12071 (*Ecballium elaterium*), MCEI-I P10296 (*Momordica charantia*), CMTI-I P01074 (*Cucurbita maxima*), MRTI-I P17680 (*Momordica repens*), CMCTI-I, CMCT-II und CMCTI-III P32041 (*Cucumis melo var. conomon*), CMTI-IV und CMTI-III P07853 (*Cucurbita maxima*), TRYPSIN INHIBITOR A P30709 (*Momordica charantia*), LLDTI-I OR LLDTI-II oder LLTI-III P26771 (*Lagenaria siceraria*), LCTI-I P25849 (*Luffa cylindrica*), BDTI-II P11968 (*Bryonia dioica*), LCTI-II P25850 (*Luffa cylindrica*), CPTI-III und CPTI-II (*Cucurbita pepo*), CSTI-IV P10292 (*Cucumis sativus*), CVTI-I P11969 (*Citrullus lanatus*), MCTI-I P10294 (*Momordica charantia*), CSTI-IIB P10291 (*Cucumis sativus*), MCTI-II P10295 (*Momordica charantia*), TGTI-I P35628 (*Luffa cylindrica*).

Hierzu wurde zunächst das Profil des jeweiligen Aminosäure-Restes von EETI-II vor dem Hintergrund der insgesamt 18 Inhibitor-Sequenzen folgendermaßen erstellt:

$$profile_{i,a_j} = \sum_{k=l}^{20} f_{i,a_k} \cdot S_{a_j a_k}$$

Hierbei ist f die Wahrscheinlichkeit, mit der der Rest a_k an Position i vorkommt und S der *score* für das Restepaar a_j,a_k. Die *scores* wurden der Blosum 62 Matrix (Henikoff & Henikoff 1992) entnommen. Hierzu ist anzumerken, dass diese *scoring*-Matrix für Proteine durchschnittlicher Größe entwickelt und ihre Anwendung daher nicht für die Analyse von konservierten Positionen in Mikroproteinen optimiert wurde. Da keine für *squahs*-Inhibitoren optimierte *scoring*-Matrix zur Verfügung stand, wurde die Blosum 62 Matrix als eine der umfassendsten Matrices für diesen Zweck ausgewählt. Um einen

ungefähren, prozentualen Wert für die Konservierung zu ermitteln, wurde der errechnete Wert für das Profil auf den *score* für jeweils übereinstimmende Reste-Paare normiert. Fehlende Positionen wurden mit dem *score* 0 belegt. Negative Profile wurden als nicht konserviert betrachtet. Auf der Grundlage der errechneten Prozentzahlen wurde die Hauptkette von EETI-II (wt) von blau (100 % Konservierung) bis rot (0 % Konservierung) koloriert (Abb. 34):

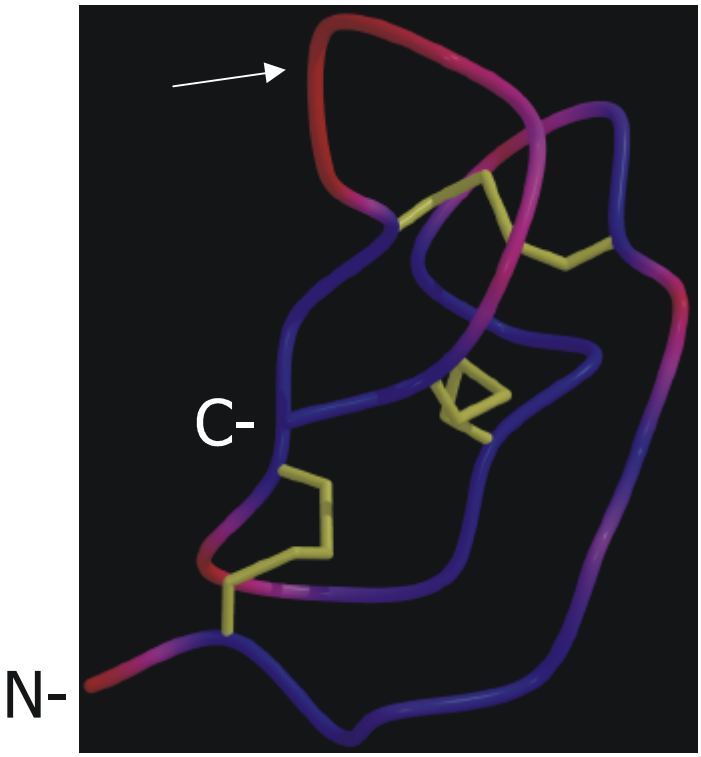


Abb. 34: Der Konservierungsgrad der Aminosäurereste in EETI-II wurde vor dem Hintergrund von 17 weiteren *squash*-Inhibitoren aus der Swiss-Prot Datenbank qualitativ berechnet. Eine blaue Farbgebung zeigt eine hohe Konservierung, eine rote Färbung eine hohe Variabilität für den entsprechenden Aminosäurerest an. Links unten im Bild liegt der n-Terminus, der GPNG-*b-turn* von EETI-II ist durch einen weißen Pfeil markiert.

Unter dem Aspekt der Funktion dieser Inhibitoren entspricht die Tatsache, dass der Bereich der Trypsin-Bindeschleife in einigen Bereichen 100% Konservierung aufweist, den Erwartungen. Allerdings sind auch die übrigen Abschnitte des Polypeptidrückgrats nach den Kriterien, die der Blosum-62-Matrix zu Grunde liegen, relativ hoch konserviert – mit Ausnahme des terminalen GPNG-*b-turns*, der über einen weiten Sequenzraum variiert ist. Diese letzte Schleife scheint demnach der einzige Bereich in den Peptidsequenzen der Inhibitor-Cystinknoten-Proteine aus den Kürbisgewächsen zu

sein, der keinem hohen evolutiven Druck unterworfen ist. Die Inhibitionskonstanten der in dieser Arbeit untersuchten EETI-II Varianten liegen im Rahmen der Messgenauigkeit im gleichen Größenbereich. Alleine die Variante $EETI^{ERDE}$ setzt sich durch einen geringfügig schlechteren K_i-Wert von den anderen untersuchten Inhibitoren ab (Wentzel *et al.* 1999). Aus den Strukturdaten ist allerdings keine signifikante Abweichung im Bereich der Protease-Inihibitorschleife ableitbar.

5.1 Zur Funktion des 22-24 *b-turns*

Der Typ I *b-turn* in EETI-II spielt, wie bereits erwähnt, ein wichtige Rolle bei der korrekten Faltung des Inhibitors. Für *squash*-Inhibitoren sind weder eine Pro-Sequenz wie bei BPTI (Weissman & Kim 1992) noch ein Vorläuferprotein wie bei den ω-Conotoxinen (Price-Carter *et al.* 1996) bekannt, aus dem das native Mikroprotein hervorgeht. Sehr wahrscheinlich existiert eine Signalsequenz zur Überführung der Polypeptidkette in das Endoplasmatische Reticulum. Die Faltung von EETI-II (wt) erfolgt *in vitro* spontan und nahezu quantitativ, auch bei der Produktion *in vivo* in *E. coli* Zellen (Christmann *et al.* 1999) verläuft die Faltung von EETI-II als Fusionspartner einer Reihe anderer Proteine korrekt. Auch in *E. coli dsba* Stämmen, die keine Disulfid-Isomerasen (Übersicht bei Novia 1999) exprimieren, bildet EETI-II die korrekte Tertiärstruktur aus.

Zur Faltungskinetik anderer *squash*-Inhibitoren waren zum Zeitpunkt dieser Arbeit keine genaueren Daten publiziert. Allerdings scheint ihre Faltung in ähnlichen Zeitdimensionen wie im Falle von EETI-II (wt) zu verlaufen. Der Zeitraum, den die Faltung von CMTI-I (Kojima *et al.* 1996) oder MCTI-A (Zhu *et al.* 2000) benötigt, liegt unterhalb eines halben Tages. Die Zeitdauer für die Faltung der drei EETI-Varianten EETI-II (wt), $EETI^{NEDE}$ und $EETI^{TNNK}$ ist mit den Summen der Positionspotentiale für den Typ I *b-turn* nach Hutchinson & Thornton (1994) korreliert. In dieser Arbeit wurden *b-turns* nach Typen klassifiziert, für jeden Typ wurde für die einzelnen Positionen der Aminosäurereste ein Potential für die Ausformung des betreffenden *b-turns* auf Grundlage statistischer Überlegungen bestimmt.

Die Rückfaltung von EETI-II (wt) verläuft am schnellsten, die aus den Positionspotentialen (PP) der Sequenz GPNG berechnete Summe für den Typ I *b-turn*

beträgt 9.2. Die Variante EETI^{NEDE} faltet mit einer PP-Summe von 7.35 bedeutend langsamer als der Wildtyp, allerdings schneller als die Variante EETI^{TNNK} mit einer PP-Summe von 5.15 für den Typ I *b-turn*. Aus den Positionspotentialen für CMTI-I mit der *b-turn* Sequenz (LEHG) errechnet sich eine PP-Summe von 5.59.

Ein Experiment, das weitere Erkenntnisse über die Funktion des #22-#24 *b-turns* bei der Faltung von *squash*-Inhibitoren liefern würde, wäre der gezielte Austausch der GPNG-Sequenz von EETI-II (wt) gegen die Aminosäuren LEHG von CMTI-I und eine kinetische Untersuchung der Faltung dieser Variante. Die zusätzliche Ermittlung der Faltungskinetik von CMTI-I würde zeigen, inwieweit die vier Aminosäuren der Typ-I *b-turn* bildenden Sequenz alleine an der Faltungsgeschwindigkeit beteiligt sind. Eine weitere, unter kinetischen Gesichtspunkten interessante Mutation wäre der Austausch von Gly #22 in EETI-II (wt) gegen ein Asparagin, das mit 2.51 ein wesentlich höheres Positions-Potential als Glycin mit 1.07 hat. Die Variante EETI^{NPNG} könnte den Erwartungen entsprechend schneller falten als EETI-II (wt). Daneben könnte untersucht werden, inwieweit Insertionen oder Deletionen von Aminosäure-Positionen im Bereich des *b-turns* die Faltung beeinflussen.

Der Bereich zwischen den beiden letzten Cysteinen weist die höchste zusammenhängende Variabilität im Vergleich zwischen EETI-II und den 17 anderen *squash*-Inhibitoren aus der Swiss-Prot Datenbank auf. Die Länge dieser Schleife kann vier Aminosäuren (z. B. MCEI-I), fünf Aminosäuren (z. B. EETI-II) oder sechs Aminosäuren (z. B. CSTI-IIB) betragen. Sollten weitere Experimente deutliche Unterschiede zwischen den einzelnen Inhibitoren in der Faltungskinetik aufzeigen, könnte das mehrere Schlüsse zulassen. Beispielsweise müsste eine schnelle Faltung nicht zwingend ein evolutives Selektionskriterium sein, die genetische Variabilität könnte daher in diesem *turn* relativ hoch sein. Dafür spricht, dass die meisten Mutationen durch einen einzigen Basenaustausch auf DNS-Ebene ineinander überführbar sind (Leluk 2000). Allerdings finden sich Hinweise darauf, dass die genaue Funktion der *squash*-Inhibitoren bisher noch nicht vollständig aufgeklärt ist. Bisher ist nur die hemmende Wirkung auf die Trypsinaktivität bekannt. Die Lokalisierung dieser Mikroproteine in Samen von Kürbisgewächsen legt die Vermutung nahe, dass die Hauptrolle der Inhibitoren in der Vorbeugung des Verdauens im Darm von Tieren nach dem Verzehr der Frucht liegt, wodurch die Samen wieder unbeschädigt ausgeschieden

werden. Die Tiere, die die Früchte zuvor gefressen haben, werden so als Vehikel für die Verbreitung der Samen benutzt, die dann fern der Mutterpflanze keimen können.

Einige dieser Pflanzen werden in anderen Kulturen als Heilpflanzen eingesetzt, beispielsweise die Bittermelone *Momordica charantia*, aus der MCTI-A isoliert wurde (Zhu *et al.* 1999). Eventuell könnten einige der *squash*-Inhibitoren Träger von weiteren Funktionen sein, die zu dieser Heilwirkung beitragen. Es ist bekannt, dass *b-turns* bei der molekularen Erkennung eine Rolle spielen können (Hutchinson & Thornton 1994). Die große Sequenzenvariabilität im Bereich zwischen den letzten beiden Cysteinen könnte ein Hinweis darauf sein, dass auf diesem Abschnitt der Struktur weitere, voneinander unterschiedliche Funktionen liegen könnten.

Mit der ständig wachsenden Anzahl von vollständig sequenzierten Genomen wird auch die zur Verfügung stehende Menge an ICK-Proteinsequenzen immer umfangreicher werden. Möglicherweise werden mit Methoden der Bioinformatik weitere Wirkungsgebiete dieser Inhibitoren erkannt, die bisher nicht in Betracht gezogen wurden.

6 Zusammenfassung

Im Rahmen des beschriebenen Promotionsprojektes sollte der *squash*-Inhibitor *Ecballium elaterium* Trypsin Inhibitor-II (EETI-II) als typischer Vertreter eines Proteins aus der Inhibitor-Cystinknoten-Familie (ICK) für die Röntgenstrukturanalyse erschlossen werden. Die nur ca. 30-50 Aminosäuren umfassenden ICK-Proteine weisen als charakteristische Strukturmerkmale neben einem dreisträngigen, antiparallelen β-Faltblatt sechs typisch verknüpfte, hoch konservierte Cysteinreste auf, wobei das erste Cystein mit dem vierten, das zweite mit dem fünften und das dritte mit dem sechsten Cysteinrest eine Disulfidbrücke ausbildet.

Das nur 36 Aminosäurereste umfassende rekombinante Protein EETI-II (wt) wurde auf DNA-Ebene gegenüber dem nativen Inhibitor leicht modifiziert und als Fusionsprotein mit einem genetisch veränderten β-Galaktosidasefragment als Einschlusskörper *in E. coli* produziert. Das um sechs c-terminale Histidine erweiterte Polypeptid wurde nach der Bromzyanspaltung über *IMAC* und *RP-HPLC* gereinigt und *in vitro* gefaltet. Für EETI-II (wt) konnten Kristallisationsbedingungen in der Raumgruppe *C2* etabliert werden, insgesamt wurden drei Synchrotron-Datensätze im Auflösungsbereich zwischen 1.6 und 1.9 Å gemessen. Ansätze zur Strukturlösung mit *MR*-Methoden blieben erfolglos. Ein Grund hierfür war, dass EETI-II (wt) voraussichtlich mit zwölf Molekülen in der asymmetrischen Einheit kristallisiert. Die Bestimmung der Schwefelatom-Positionen des Cystinknotens von EETI-II (wt) über das anomale Signal führte ebenfalls nicht zur Strukturlösung, da die Streukraft der Kristalle nicht die für dieses Experiment benötigte Datenqualität lieferte. Mit EETI-II (wt) Kristallen konnte ein Jodid-Derivat generiert werden, das in weiteren Experimenten zur Lösung des Phasenproblems eingesetzt werden soll.

EETI-II (wt) konnte im Komplex mit Trypsin aus dem Schweinepankreas in der Raumgruppe $P4_132$ kristallisiert werden, die Struktur wurde gegen Synchrotron-Daten bis 1.5 Å Auflösung bis zu einem *R*-Wert von 21.4 % verfeinert. Es konnte gezeigt werden, dass EETI-II (wt) die typischen Strukturmerkmale eines Cystinknoten-Proteins aus der Gruppe der *squash*-Inhibitoren besitzt. Der *b-turn* der Positionen #22 bis #25 konnte als Typ I *b-turn* klassifiziert werden.

Von der EETI-II *b-turn*-Varianten EETI^{TNNK}, die von der Arbeitsgruppe Dr. Kolmar aus einem zufallsmässig erzeugten Repertoire unter dem Kriterium der Trypsinbindung selektiert worden war, und deren *in vitro* Faltungskinetik stark von der der Varianten EETI-II (wt) abweicht, konnte eine Struktur in der Raumgruppe *I4* bei einer Auflösung von 1.9 Å bis zu einem *R*-Wert von 13.6 % verfeinert werden. Von zwei anderen *b-turn*-Varianten wurden die Strukturen im Anfangsstadium bestimmt. Es konnte gezeigt werden, dass sich die faltungskinetisch sehr einflussreichen Mutationen im #22 bis #24 *b-turn* von EETI-II nicht in der Konformation des Hauptkettenverlaufs der einzelnen EETI-II-Varianten reflektieren. Die #22 bis #24 *b-turns* der EETI-II-Varianten waren untereinander praktisch identisch, zudem waren sie mit einer mittleren Koordinaten-Abweichung von unter 0.3 Å in den Cα-Atompositionen zum entsprechenden *b-turn* des *squash*-Inhibitors CMTI-I superpositionierbar. In weiteren Experimenten könnten die strukturellen und faltungskinetischen Folgen von Insertionen, Deletionen und weiteren Aminosäuresubstitutionen im Sequenzbereich #22 bis #25 von EETI-II (wt) untersucht werden.

7 Literatur

Amemiya, Y., Satow, Y., Matsushita, T., Chikawa, J., Wakabayashi, K., Miyahara, J. (1988): A Storage Phosphor Detector (Imaging Plate) and its Application to Diffraction Studies Using Synchrotron Radiation. *Topics in Current Chemistry*, **147**: 121-144

Bairoch, A., Apweiler, R. (2000): The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **28**: 45-48

Beliakova, A., Schumakovitch, A. (2000). Auflösung mathematischer Knoten. *Uni Nova (Wissenschaftsmagazin der Universität Basel)* **87**: 45-50
URL: http://www.zuv.unibas.ch/uni_nova/087/toc.shtml

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. Shindyalov, I. N., Bourne, P. E. (2000): The Protein Data Bank. *Nucleic Acids Research* **28**: 235-242

Biou, V., Dumas, R., Cohen-Addad, C., Douce, R., Job, D., Pebay-Peyroula, E. (1997): The crystal structure of plant acetohydroxy acid isomeroreductase complexed with NADPH, two magnesium ions and a herbicidal transition state analog determined at 1.65 Å resolution. *EMBO J.* **16**: 3405-15

BobScript: V. 2.5 (1994-1999) © ? Robert Esnouf.

Bolewska, K., Krowarsch, D., Otlewski, J., Jaroszewski, L., Bierzynski, A. (1995): Synthesis, cloning and expression in Escherichia coli of a gene coding for the Met8-->Leu CMTI I--a representative of the squash inhibitors of serine proteinases. *FEBS Lett.* **377**: 172-4

Brünger, A. T. (1993): Assessment of phase accucy by cross validation: The free R value.
Methods and applications. *Acta Cryst. D* **49**: 24-36.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., Warren, G. L.(1998): Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**: 905-21

Burmeister, W. P. (2000): Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. D* **56**: 328-41

Christmann, A., Walter, K., Wentzel, A., Krätzner, R., Kolmar, H. (1999): The cystine knot of a squash-type protease inhibitor as a structural scaffold for Escherichia coli cell surface display of conformationally constrained peptides. *Protein Eng.* **12**: 797-806

Chiche, L., Gaboriaud, C., Heitz, A., Mornon, J. P., Castro, B., Kollman, P.A. (1989): Use of restrained molecular dynamics in water to determine three-dimensional protein structure: prediction of the three-dimensional structure of Ecballium elaterium trypsin inhibitor II. *Proteins* **6**: 405-17

Daly, N. L., Craik, D. J. (2000): Acyclic permutants of naturally occurring cyclic proteins. Characterization of cystine knot and beta-sheet formation in the macrocyclic polypeptide kalata B1. *J. Biol. Chem.* **275**: 19068-75

Dauter, Z.(1999): Data-collection strategies. *Acta Crystallogr. D* **55**: 1703-17

Dauter, Z., Dauter, M. (2001): Entering a New Phase. Using Solvent Halide Ions in Protein Structure Determination. *Structure* **9**: R21-R26

Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G., Sheldrick, G. M. (1999): Can Anomalous Signal of Sulfur Become a Tool for Solving Protein Crystal Structures? *J. Mol. Biol.* **289**: 83-92

Dower, W. J., Miller, J. F., Ragsdale, C. W. (1988): High eficiency transformation of E. Coli by high voltage elektroporation, *Nucl. Acid. Res.* **16**: 6127-6145

Drenth, J. (1994): Principles of Protein X-Ray Crystallography, 1st Edition
Springer Verlag, New York

Ducruix, A., Griegé, R. (1992): Crystallization of Nuclic Acids and Proteins. 1. Aufl., Oxford University Press, New York 73-98.

Esnouf, R. M.(1999): Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. Acta Crystallogr D **55**: 938-40

Favel, A., Mattras, H., Coletti-Previero, M. A., Zwilling, R., Robinson, E. A., Castro, B. (1989): Protease inhibitors from Ecballium elaterium seeds. *Int. J. Pept. Protein. Res.* **33**: 202-8

Fletcher, J. I., Dingley, A. J., Smith, R., Connor, M., Christie, M. J., King, G. F.(1999): High-resolution solution structure of gurmarin, a sweet-taste-suppressing plant polypeptide. *Eur. J. Biochem.* **264**: 525-33

FRAMBO: Programm zur Diffraktometersteuerung und Datensammlung mit
Flächenzählern. (1994), © Siemens.

Garman, E.(1999): Cool data: quantity AND quality. *Acta Crystallogr. D* **55**: 1641-53

Gruner, S. M., Ealick, S. E. (1995): Charge coupled device X-ray detectors for macromolecular crystallography. *Structure* **3**:13-5

Gursky, O., Badger, J., Li, Y., Caspar D. L. (1992): Conformational changes in cubic insulin crystals in the pH range 7-11. *Biophys. J.* **63**: 1210-20

Guddat, L. W., Martin, J. A., Shan, L., Edmundson, A. B., Gray, W. R. (1996): Three-dimensional structure of the alpha-conotoxin GI at 1.2 Å resolution. *Biochemistry* **35**: 11329-35

Hendrickson, W. A., Teeter, M. M. (1981): Structure of the hydrophobic protein crambin, determined directly from the anomalous scattering of sulfur. *Nature* **290**: 107-113

Heitz, A., Chiche, L., Le-Nguyen, D., Castro, B. (1989): 1H 2D NMR and distance geometry study of the folding of *Ecballium elaterium* trypsin inhibitor, a member of the squash inhibitors family. *Biochemistry* **28**: 2392-8

Heitz, A., Le-Nguyen, D., Chiche, L. (1999): Min-21 and min-23, the smallest peptides that fold like a cystine-stabilized beta-sheet motif: design, solution structure, and thermal stability. *Biochemistry* **38**: 10615-25

Henikoff, S., Henikoff, J. G. (1992): Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci.* **89**: 10915-10919

Helland, R., Berglund, G. I., Otlewski, J., Apostoluk, W., Andersen, O. A., Willassen, N. P., Smalas, A O. (1999): High-resolution structures of three new trypsin-squash-inhibitor complexes: a detailed comparison with other trypsins and their complexes. *Acta. Crystallogr. D* **55**:139-48

Helliwell, J. R., (1992): Macromolecular Crystallography with Synchrotron Radiation
Cambridge University Press.

HKL2000. Copyright © 2000 HKL Research, Inc.

Huang, Q., Liu, S., Tang, Y. (1993): Refined 1.6 Å resolution crystal structure of the complex formed between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the squash family. Detailed comparison with bovine beta-trypsin and its complex. *J. Mol. Biol.* **229**: 1022-36

Hutchinson, E. G., Thornton, J. M. (1994): A revised set of potentials for beta-turn formation in proteins. *Protein Sci.* **3**: 2207-16

Hodgkin, D. C. (1974): Proceedings: Varieties of insulin. The Sir Henry Dale lecture for 1974. *J Endocrinol.* **63**: 3P-14P

Jacraik, J. & Kin, S.-H. (1991): Sparse matrix sampling : a screening method for crystallization of proteins. *J. Appl. Cryst.* **24**: 409-411

Kagawa, Y. (1999): Biophysical studies on ATP synthase. *Adv. Biophys.* **36**: 1-25

Kissinger, C. R., Gehlhaar, D. K., Fogel, D. B. (1999): Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr. D* **55**: 484-91

Klebe, G. (2000): Recent developments in structure-based drug design. *J. Mol. Med.* **78**: 269-81

Kohno, T., Kim, J. I., Kobayashi, K., Kodera, Y., Maeda, T., Sato, K. (1995): Three-dimensional structure in solution of the calcium channel blocker omega-conotoxin MVIIA. *Biochemistry* **34**: 10256-65

Kojima, S, Miyoshi, K., Miura, K.(1996): Synthesis of a squash-type protease inhibitor by gene engineering and effects of replacements of conserved hydrophobic amino acid residues on its inhibitory activity. *Protein Eng.* **9**: 1241-6

Lattman, E. (1985): Diffraction methods for biological macromolecules. Use of the rotation and translation functions. *Methods Enzymol.* **115**: 55-77

Laskowski, M., Jr., Kato, I. (1980): Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**: 593-626

Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, M. T. (1993): PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**: 283-291

Leluk, J. (2000): Serine Proteinase Inhibitor Family in Squash Seeds: Mutational Variability Mechanism and Correlation. *Cellular & Molecular Biology Letters* **5**: 91-106

Le-Nguyen, D., Heitz, A., Chiche, L., el Hajji, M., Castro, B. (1993): Characterization and 2D NMR study of the stable [9-21, 15-27] 2 disulfide intermediate in the folding of the 3 disulfide trypsin inhibitor EETI-II. *Protein Sci.* **2**: 165-74

Le-Nguyen, D., Nalis, D., Castro, B. (1989): Solid phase synthesis of a trypsin inhibitor isolated from the Cucurbitaceae Ecballium elaterium. *Int. J. Pept. Protein Res.* **34**: 492-7

Lewis, R. (1994): Multiwire Gas Proportional Counters: Decrepit Antiques or Classic Performers? *J. Synchrotron Rad.* **1**: 43-45

McPherson, A. (1998): Crystallisation of Biological Macromolecules. *J. Crystal Growth* **122**: 161-167

McRee, D. (1999): XtalView/Xfit – A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *J.Struct.Biol.* **125**: 156-165

Matthews, B. W. (1968): Solvent content of protein crystals. *J. Mol. Biol.* **33**: 491-7

Merritt, E. A., Bacon, D. J. (1997): Raster 3D: Photorealistic Molecula graphics, *Meth.Enzymol.* **277**: 505-524.

Miller, R., DeTitta, G.T., Jones, R., Langs, D.A., Weeks, C.M. Hauptman, H.A. (1993): On the application of the minimal principle to solve unknown structures. *Science* **259**: 1430-1433.

Murshudov, G. N., Vagin A. A., Dodson, E. J. (1997): Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst. D* **53**: 240-255.

Navaza, J. (1994): AMoRe: an automated package for molecular replacement. *Acta Cryst. A* **50**: 157-163

Nilges, M., Habazettl, J., Brunger, A. T., Holak, T. A. (1991): Relaxation matrix refinement of the solution structure of squash trypsin inhibitor. *J. Mol. Biol.* **219**: 499-510

Ninomiya, Y., Imoto, T. (1995): Gurmarin inhibition of sweet taste responses in mice. *Am. J. Physiol.* **268**: R1019-25

Norledge, B. V., Hay, R. E., Bateman, O. A., Slingsby, C., Driessen, H. P. (1997): Towards a molecular understanding of phase separation in the lens: a comparison of the X-ray structures of two high Tc gamma-crystallins, gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and gammaD. *Exp Eye Res* **65**: 609-30

Norton, R. S., Pallaghy, P. K. (1998): The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* **36**:1573-83

Noiva, R. (1999): Protein disulfide isomerase: the multifunctional redox chaperone of the endoplasmic reticulum. *Semin. Cell Dev. Biol.* **10**: 481-93

O'Keeffe, M., Brese, N.E. (1991): Atom Sizes and Bond Lengths in Molecules and Crystals. *J. Am. Chem. Soc.* **113**: 3226-3229

Olivera, B. M., Gray, W. R., Zeikus, R., McIntosh, J. M., Varga, J., Rivier, J., de-Santos, V., Cruz, L. J. (1985): Peptide neurotoxins from fish-hunting cone snails. *Science* **230** (4732): 1338-43

Olivera, B. M., Hillyard, D. R., Marsh, M., Yoshikami, D. (1995): Combinatorial peptide libraries in drug design: lessons from venomous cone snails. *Trends. Biotechnol.* **13**: 422-6

Olivera, B. M. McIntosh, J. M. Cruz, L. J. Luque, F. A.; Gray, W. R. (1984): Purification and sequence of a presynaptic peptide toxin from Conus geographus venom. *Biochemistry*. **23**: 5087-90.

Otlewski, J., Zbyryt, T., Krokoszynska, I., Wilusz, T., (1990): Inhibition of serine proteinases by squash inhibitors. *Biol. Chem. Hoppe. Seyler*. **371**: 589-94.

Otlewski, J., Zbyryt, T. (1994): Single peptide bond hydrolysis/resynthesis in squash inhibitors of serine proteinases. 1. Kinetics and thermodynamics of the interaction between squash inhibitors and bovine beta-trypsin. *Biochemistry* **33**: 200-7

Ottani, V., Raspanti, M., Ruggeri, A. (2001): Collagen structure and functional implications. *Micron* **32**:251-60

Otwinowski Z., Minor W. (1997): Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Meth. Enzymol.* **276**: 307-326.

Pape, T., Schneider, T. R., Sheldrick, G. M., Kuhn, P., Graeme Laver, W., Vasella, A., Vorwerk, S., Vimr, E., Garman, E. : *to be published*

Porath, J., Olin, B., Granstrand, B. (1983): Immobilized-metal affinity chromatography of serum proteins on gel-immobilized group III A metal ions. *Arch. Biochem. Biophys.* **225**: 543-7

Price-Carter, M., Gray, W. R., Goldenberg, D. P.(1996): Folding of omega-conotoxins. 2. Influence of precursor sequences and protein disulfide isomerase. *Biochemistry* **35**: 15547-57

Ramachandran, G. N., Sassiekharan, V. (1968): Conformation of polypeptides and proteins. *Adv. Protein Chem.* **28**: 283-437.

Ravelli, R. B. G., Sweet, R. M., Skinner, J. M., Duisenberg, A. J. M. Kroon, J. (1997): STRATEGY: a program to optimize the starting spindle angle and scan range for X-ray data collection. *J. Appl. Cryst.* **30:** 551-54

Read, R. .J. (1986): Improved Fourier Coefficients for Maps Using Phases from Partial Structures with Errors. *Acta Cryst. A* **42**: 120-129.

Rossmann, M. G. (1990): The molecular replacement method. *Acta Crystallogr A* **46**: 73-82

SAINT: Programm zur Integration und Reduktion von kristallographischen Flächenzählerdaten. V 5.00 (1997), Bruker Analytical X-ray systems, © Bruker AXS.

Savarin, P., Guenneugues, M., Gilquin, B., Lamthanh, H., Gasparini, S., Zinn-Justin, S., Menez, A. (1998): Three-dimensional structure of kappa-conotoxin PVIIA, a novel potassium channel-blocking toxin from cone snails. *Biochemistry* **37**: 5407-16

Schäfer, M., Schneider, T. R., Sheldrick, G. M. (1996): Crystal structure of vancomycin. *Structure*. **4**: 1509-15

Schägger, H., von Jagow, G. (1987): Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.***166**: 368-79

Schechter, I., Berger, A. (1967): On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**: 157-62

Sheldrick, G. M. (1997): SHELX applications to macromolecules. In "Direct methods for solving macromolecular structures". S. Fortier, Ed., Kluwer Academic Publishers, pp. 401-411.

Sheldrick, G. M. (1997): SADABS, Programm zur Absorptionskorrektur von Bruker (Siemens) Flächenzählerdaten. Universität Göttingen.

Scheufler, C., Sebald, W., Hulsmeyer, M. (1999): Crystal structure of human bone morphogenetic protein-2 at 2.7 Å resolution. *J. Mol. Biol.* **287**: 103-15

Shon, K. J., Hasson, A., Spira, M. E., Cruz, L. J., Gray, W. R., Olivera, B. M (1994): Delta-conotoxin GmVIA, a novel peptide from the venom of Conus gloriamaris. *Biochemistry*. **33**: 11420-5

Shon, K. J., Grilley, M. M., Marsh, M., Yoshikami, D.; Hall, A. R., Kurz, B., Gray, W. R.; Imperial, J. S., Hillyard, D. R., Olivera, B. M, (1995): Purification, characterization, synthesis, and cloning of the lockjaw peptide from Conus purpurascens venom. *Biochemistry*. **18**: 4913-8.

Stubbs, M. T., Morenweiser, R., Sturzebecher, J., Bauer, M., Bode, W., Huber, R., Piechottka, G. P., Matschiner, G., Sommerhoff, C. P., Fritz, H., Auerswald, E. A. (1997): The three-dimensional structure of recombinant leech-derived tryptase inhibitor in complex with trypsin. Implications for the structure of human mast cell tryptase and its inhibition. *J. Biol. Chem.* **272**: 19931-7

Taylor, W. R.(2000): A deeply knotted protein structure and how it might fold. *Nature* **406**: 916-9

Temple, M. D., Hinds, M. G., Sheumack, D. D., Howden, M. E., Norton, R. S. (1999): 1H NMR study of robustoxin, the lethal neurotoxin from the funnel web spider *Atrax robustus*. *Toxicon* **37**: 485-506

Usón, I., Sheldrick, G. M. (1999): Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9**: 643-8

Usón, I., Sheldrick, G. M., de La Fortelle, E., Bricogne, G., Di Marco, S., Priestle, J. P., Grutter, M. G., Mittl, P. R. (1999): The 1.2 Å crystal structure of hirustasin reveals the intrinsic flexibility of a family of highly disulphide-bridged inhibitors. *Structure Fold. Des.* **7**: 55-63

Vriend, G. (1990): WHAT IF: a molecular modelling and drug design program, *J. Mol. Graph.* **8**: 52-56

Wakamatsu, K., Kohda, D., Hatanaka, H., Lancelin, J. M., Ishida, Y., Oya, M., Nakamura, H., Inagaki, F., Sato, K. (1992): Structure-activity relationships of mu-conotoxin GIIIA: structure determination of active and inactive sodium channel blocker peptides by NMR and simulated annealing calculations. *Biochemistry* **31**: 12577-84

Weik,. M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J., Sussman, J. L. (2000): Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc. Natl. Acad. Sci.* **97**: 623-8

Weissman, J. S., Kim, P. S. (1992): The pro region of BPTI facilitates folding. *Cell* **71**: 841-51

Wentzel, A., Christmann, A., Krätzner, R., Kolmar, H. (1999): Sequence requirements of the GPNG beta-turn of the Ecballium elaterium trypsin inhibitor II explored by combinatorial library screening. *J. Biol. Chem.* **274**: 21037-43

XPREP: Programm zur Bearbeitung von Beugungsdaten & Untersuchung reziproker Gitter. V. 5.10 (1997) – 6.09 (2001), © ? Bruker AXS.

Zhu, Y., Huang, Q., Qian, M., Jia, Y., Tang, Y. (1999): Crystal structure of the complex formed between bovine beta-trypsin and MCTI-A, a trypsin inhibitor of squash family, at 1.8- Å resolution. *J. Protein. Chem.* **18**: 505-9

Publikationen

Walawalkar, M. G., Murugavel, R., Roesky, H. W., Usón, I. and Krätzner, R. (1998): Gallophosphonates Containing Alkali Metal Ions. 2.Synthesis of Gallophosphonates Incorporating Na+ and K+ Ions. *Inorg. Chem.* **37**: 473 -478

Chakraborty, D., Bhattacharjee, M., Krätzner, R., Siefken, R.,Roesky, H. W., Usón, I. And Schmidt, H. G. (1999): First Structurally Characterized Organometallic Chloro oxo-Peroxo Compounds of Molybdenum and Tungsten. *Organometallics* **18**: 106-108

Wentzel, A., Christmann, A., Krätzner, R., Kolmar, H. (1999):Sequence requirements of the GPNG ß-turn of the Ecballium elaterium trypsin inhibitor II explored by combinatorial library screening. *J. Biol. Chem.* **274**: 21037-43

Christmann, A., Walter, K., Wentzel, A., Krätzner, R., Kolmar, H. (1999): The cystine knot of a squash-type protease inhibitor as a structural scaffold for Escherichia coli cell surface display of conformationally constrained peptides. *Protein Eng.* **12**: 797-806

Mösch-Zanetti, N., Krätzner, R., Lehmann, C., Schneider, T., R., Usón, I. (2000): Titan (III) Compunds with hapto-2-pyrazolato Ligands. *Eur. J. Inorg. Chem.* **13-16** : 1549-1555

Chakraborty, C., Chandrasekhar, V., Bhattacharjee, M., Krätzner, R., Roesky, H. W., Noltemeyer, M. and Schmidt, H.-G. (2000): Metal Alkoxides as Versatile Precursors for Group 4 Phophates: Synthesis and X-ray Structure of a Novel Organosoluble Zirconium Phophonate. *Inorg. Chem.* **39**: 23-26.

C. Ackerhans, B. Räke, R. Krätzner, P. Müller, H. W. Roesky, I. Usón (2000): Ammonolysis of Trichlorsilanes. *Europ. J. Inorg. Chem.* **5**: 827-830.

Danksagung

Herrn Professor Sheldrick möchte ich für die interessante Aufgabenstellung meines Promotionsthemas, für die konstruktive Unterstützung meiner Arbeit und für seine ständige Diskussionsbereitschaft danken.

Die im Rahmen dieses Projekts notwendigen molekularbiologischen Arbeitsschritte wurden in Laborräumen und mit Hilfe von Geräten des „Instituts für Molekulare Genetik und präparative Molekularbiologie“ der Universität durchgeführt. Herrn Prof. Dr. Fritz möchte an dieser Stelle für die hervorragenden Arbeits- und Experimentiermöglichkeiten danken, die mir wie einem Abteilungsmitglied zur Verfügung standen.

Frau Dr. Isabel Usón danke ich für die engagierte Betreuung dieses Projekts, Ihre ständige Gesprächsbereitschaft und für die Vermittlung kristallographischen Wissens.

Herrn PD Dr. Kolmar danke ich für die hervorragende Zusammenarbeit bei der Etablierung der Aufreinigung von EETI-II Varianten. Herrn Dr. Thomas Schneider danke ich für die vielen anregenden Diskussionen über Kristallographie, das Messen des EETI-II (wt) Datensatzes auf der ESRF-*beamline* und seine Unterstützung bei Messungen am DESY. Herrn Niklas Peinecke danke ich für die Erklärungen zur Knotentheorie, Fabio Dall’Antonia, Peter Müller, Thomas Pape, Stina Schiller und Alexander Wentzel für das Korrekturlesen dieser Arbeit.

Allen Mitarbeitern meiner Arbeitsgruppe und des „Instituts für Molekulare Genetik und Präparative Molekularbiologie“ danke ich für das hervorragende Arbeitsklima, das wesentlich zum Gelingen dieser Arbeit beigetragen hat.

Meinen Eltern möchte ich unter anderem dafür danken, dass sie mir ein sorgenfreies Studium ermöglicht haben. Meinen Großeltern und Frau Liselotte Degler danke ich dafür, dass sie mich überall dort unterstützt haben, wo sie es konnten.

Lebenslauf

Am 9. September 1971 wurde ich als Kind von Günter und Ursula Krätzner, geb. Rieth, in Fulda geboren. Ich besuchte die Winfriedschule Fulda und bestand am 10. Juni 1991 die Abiturprüfung.

Im Wintersemester 1991/1992 immatrikulierte ich mich an der Georg-August-Universität Göttingen für den Studiengang Biologie/Diplom. Im Wintersemester 1993/1994 immatrikulierte ich mich an der gleichen Universität zusätzlich für den Studiengang Chemie auf Diplom.

Am 22. Oktober 1993 bestand ich die Vordiplomsprüfung im Fach Biologie, am 30. Juni 1995 absolvierte ich die Vordiplomsprüfung im Fach Chemie. Im Wintersemester 1996/1997 begann ich mit den Experimenten zu meiner Diplomarbeit in der Arbeitsgruppe von Herrn Professor Dr. Fritz am „Institut für molekulare Genetik“ der Universität zum Thema „Zur Anwendung des *Ecballium elaterium* Trypsin Inhibitors (EETI-II) als Strukturgerüst für die Präsentation eines konformationell eingeschränkten Peptidrepertoires“. Am 19. März 1997 bestand ich die Diplomprüfung, seit April 1997 arbeite ich in der Arbeitsgruppe von Professor Sheldrick, Ph.D., am Institut für Anorganische Chemie der Universität an meiner Promotion zum Thema: „Strukturelle Untersuchungen an Varianten des *Ecballium elaterium* Trypsin Inhibitors-II (EETI-II)“ in Zusammenarbeit mit der Arbeitsgruppe von Herrn PD Dr. Kolmar am „Institut für molekulare Genetik und präparative Molekularbiologie“ der Universität.