Non-Enzymatic Extracellular Proteins and Their Role in Growth, Defense, and Metabolism

Matthew Daniel McDougall

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

University of Manitoba

Winnipeg

Abstract

While enzymes can be easily categorized by their catalytic function, the functions of non-enzymatic proteins can be much more subtle. In the absence of experimental data, protein functions are often assigned based on sequence or domain conservation, but this can be misleading. Beginning with the human guidance protein Netrin-4, to the pit viper venom constituent Rhodocetin and its interaction integrin α2β1, a blood clotting protein, and finishing with Right Handed Coiled-Coil, a surface layer component of microorganisms that live in the harshest environments, this thesis brings to light new knowledge on three extracellular systems where structural characterization and subsequent validation were used to clarify their functions or their mechanism of action.

Acknowledgements

As evidenced by the substantial number of authors in the papers and citations, science is not a solitary pursuit. I could not have completed this degree on my own.

I could not have asked for a better supervisor than Jörg Stetefeld; without his support, counsel, and direction my tenure as a student would have been decidedly less fulfilling. Beside Dr. Stetefeld himself is the number of immensely talented research associates, post-docs, and students, past and present, that form the Stetefeld research group. I would especially like to thank Drs. Markus Meier, Trushar Patel, and George Orriss, who all offered the guidance I needed when I first started in the lab in 2012, and continued to do so as my studies progressed.

The Department of Chemistry, both students and faculty, whose doors are always open for discussion, helped me work through a number of difficult problems and shared in my successes over the last few years. This department with its diverse interests has helped me fill some of my many gaps in knowledge.

I would be remiss if I didn’t also include the Department of Microbiology, especially Ben Bailey-Elkin from Brian Mark’s group. We both jumped into the field of macromolecular crystallography at the same time, and spent many hours arguing about the more esoteric points of the discipline.

The help of my advisory committee, Drs. Sean McKenna, Gregg Tomy, and Suresh Mishra, who have followed my work closely and have always been ready to provide another perspective and proffer advice, has always been greatly appreciated.

And lastly, I have to thank my friends and family outside of campus who have been patient and understanding of the commitments required to complete a doctoral degree. Kelly, thank you for always being there for me, even when you were thousands of kilometres away.

Table of Contents

Abstract i

Acknowledgements ii

Table of Contents iv

List of Tables xiii

List of Figures xiii

List of Abbreviations xvi

Chapter 1: Non-enzymatic Extracellular Proteins and Their Role in Growth, Defence, and Metabolism 1

- Introductory Remarks 1
- An Overview of Netrins and the Netrin Family of Proteins 2
  - Laminins and the Basement Membrane 2
  - Laminin Domain Topology 3

    A Brief Overview of the Netrins ... 4
    Netrin 4 ... 5
  Haemostasis and its Antagonism by *Calloselasma rhodostoma* ... 6
    Mammalian Haemostasis ... 6
    Venom of the Malayan Pit Viper ... 7
  The Surface Layer and Metabolism of *Staphylothermus marinus* ... 8
    Archaeal Surface Layers ... 8
    Surface Layer of *Staphylothermus marinus* ... 9
    Elemental Sulfur Metabolism in Desulfurococcales ... 10
  References ... 11
Chapter 2: Instrumentation and Methods ... 19
  Preparation of RHCC-6His *E. coli* ... 19
  Expression and Purification of RHCC ... 20
  Dynamic Light Scattering ... 21
  Circular Dichroism ... 22

X-ray Crystallography 23

Introduction to Macromolecular Crystallography 23

Preparation and Crystallization of RHCC-S_8 23

Crystallization of Netrin-4ΔC 24

Data Collection, Processing and Structure Solution of RHCC-S_8 24

Anomalous Sulfur Signal 25

Data Collection, Processing and Structure Solution of Net4ΔC 26

Data Collection, Processing and Structure Solution of Rhodocetin-γδ in complex with the Integrin α2A domain 26

References 27

Chapter 3: Structural decoding of netrin-4 reveals unique non-enzymatic disruptive forces towards mature basement membranes 31

Preface 34

Contributions of Authors 34

Abstract 35

Introduction 36

Results 39

Net4 has unique structural features absent in other netrins 39

Net4 does not interact with Net1 receptors 42

Net4 forms a high affinity complex with laminin γ1 45

Net4 disrupts pre-existing laminin networks 48

Net4 interferes with BM assembly on the cell surface 51

Net4 may affect axon outgrowth through matrix reorganization 54

Endothelial tubes require a functional basement membrane 56

Net4 inhibits tumor growth through laminin binding 62

Discussion 68

Tables 72

Methods 73

Recombinant protein expression and purification 73

Design of chimeric constructs and site-directed mutagenesis 74

Structure determination of Net4 75

Octet binding studies 75

Microscale thermophoresis binding assay 76

Laminin network disruption assay 77

Laminin assembly on Schwann cell surface 77

Neurite outgrowth assay 78

Cell culture 79

Tube-like formation assay 80

Three-dimensional culture in collagen I matrix 80

Chick Chorioallantoic Membrane angiogenesis assay 81

Immunohistochemistry of the CAM 82

Transmission electron microscopy analyses of CAM capillaries 83

Therapeutic treatment of melanoma 83

Statistics 84

References 84

Chapter 4: Dramatic and concerted conformational changes enable rhodocetin to block α2β1 integrin selectively 94

Preface 95

Author Contributions 95

Abstract 96

Introduction 97

Results 101

- Purification and characterization of the RCγδ-α2A complex 101
- Molecular structure of the rhodocetin γδ-α2A complex 103
- Rhodocetin RCγδ binds the “closed” conformation of α2A 106
- The epitope of the monoclonal antibody IIIG5 is unmasked in the RCγδ-α2A complex 109
- Conformational changes within the RCγδ-dimer after α2A binding 112
- Interaction of the RCγ subunit with loop2 of α2A-domain is essential for RC binding to the integrin 117

Discussion 120

Tables 128

Materials and Methods 129

Materials ... 129

Tryptophan-specific chemical modification of RC ... 129

Isolation of RCγδ-α2A complex ... 130

Crystallization, data processing and structure refinement ... 131

Generation of integrin α2 A-domain mutants ... 132

Binding and inhibition assays of α2A-domain with RC ... 135

Capturing ELISA with IIIG5 ... 136

Isolation of IIIG5 epitope and mass spectrometry (MS) ... 136

Mathematical evaluation of titration curves ... 138

Statistical analysis. ... 138

References ... 139

Chapter 5: Archaea S-layer nanotube from a “black smoker” in complex with cyclo-octasulfur $S8$ rings ... 150

Preface ... 151

Contributions of Authors ... 151

Abstract ... 152

Introduction 152

Results 154

Biophysical and structural elucidation of the $RHCC - S8$ complex 154

Stability of RHCC-S_8 vs apo-RHCC 159

Discussion 160

Tables 163

Methods 165

Expression and Purification of RHCC 165

DLS Experiments 166

CD Experiments 166

Crystallization, Data processing and Structure refinement 166

Molecular Dynamic simulations 167

References 173

Chapter 6: Summary and Conclusions 179

References 183

Appendix I: Additional Supporting Information for Structural decoding of netrin-4 reveals unique non-enzymatic disruptive forces towards mature basement membranes. 185

Statement of Study Approval 185

Tables 186

Figures 188

References 202

Appendix II: Additional Supporting Information for Dramatic and concerted conformational changes enable rhodocetin to block $\alpha 2\beta 1$ integrin selectively 204

Acknowledgements 204

Tables 205

Figures 207

Appendix III: Additional Supporting Information for Archaea S-layer nanotube from a “black smoker” in complex with cyclo-octasulfur S_8 rings 214

Acknowledgements 215

Figures 216

Right to Reproduce Licence 214

List of Tables

Table 3-1. Data collection and refinement statistics 72

Table 4-1. Data and refinement statistics of the RCγδ-α2A crystal structure 128

Table 5-1. Crystallographic table (Sulfur-native/ S-SAD)[a] 163

Table 5-2. Geometry of individual S_8 ring systems. 164

Table 5-3. Calculated Contributions to the Absolute Free Energy of Transfer (kJ/mol). 165

Table I-1. Hydrodynamic data for Net4-ΔC, LNγ1, Net4-ΔC-γ1LN-LEa1-4, and Net4-FL-γ1LN-LEa1-4 186

Table II-1. PCR primers used to generate integrin mutants 205

List of Figures

Figure 3-1. Overall Structure of Net4. 41

Figure 3-2. Binding comparison of Net4 and Net1 to netrin receptors and laminin. 44

Figure 3-3. Net4 disrupts pre-existing laminin network 51

Figure 3-4. Net4 blocks laminin assembly on the cell surface. 53

Figure 3-5. Net4 induces neurite outgrowth in a laminin dependent manner. 55

Figure 3-6. Net4 activity on angiogenesis is dependent on Net4/laminin interaction. 61

Figure 3-7. Laminin polymers are essential to maintain capillary networks and tumor progression. 63

Figure 3-8. Impact of Net4 on disruption of the laminin network and its effect on vascularization. 67

Figure 4-1. Isolation of the rhodocetin γδ-α2A complex on Ni Sepharose column. 102

Figure 4-2. The molecular structure of the RCγδ-α2A complex. 104

Figure 4-3. Rhodocetin recognizes the “closed” conformation, but not the “open” conformation of the integrin α2A domain. 108

Figure 4-4.The monoclonal antibody IIIG5 recognizes its epitope within the RCγ subunit in the RCγδ-α2A complex but not in the tetrameric RCαβγδ. 111

Figure 4-5. Conformational changes of RCγδ upon α2A binding. 113

Figure 4-6. An overview of the RCγ and RCδ binding residues, depicting the local conformational changes which occur upon α2A binding. 116

Figure 4-7. Loop 2 of the α2A domain is the interaction site for the RCγ subunit. 120

Figure 4-8. A comparison of the RCγδ-α2A and EMS16αβ-α2A binding interfaces. 124

Figure 4-9. Molecular mechanism of the RCγδ-α2A interaction. 126

Figure 5-1. RHCC in complex with cyclo-octasulfur S_8 crowns.. 156

Figure 5-2. Thermostability of the RHCC tetramer.. 157

Figure 5-3. Comparison of the nine-water cluster with the S_8 crown in cavity 2...................... 160

Figure I-1. Alignment of netrins and laminins.. 189

Figure I-2. Structural features of Net4-ΔC.. 190

Figure I-3. Net4 binding to integrins. 191

Figure I-4. Alignment of netrins and laminins.. 192

Figure I-5. Laminin-γ1 interferes with the Net4-FL dimer resulting in a 1:1 complex............. 193

Figure I-6. Identification of the Lmγ1 binding epitopes in Net4.. 195

Figure I-7. The laminin network within Matrigel is necessary to establish and maintain endothelial tube-like structures.. 197

Figure I-8. Inhibition of endothelial tube-like structures through Net4.................................... 199

Figure I-9. Inhibition of B16-F1 tube-like structure formation through Net4.......................... 200

Figure I-10. Inhibitory activity of Net4 is independent from matrix deposition. 201

Figure I-11. Overview of recombinant proteins used in this study. 202

Figure II-1. Asymmetric Unit of the RCγδ-α2A crystal structure. 207

Figure II-2. Molecular model of the disulfide-locked conformation mutants of α2A domain. 208

Figure II-3. Identification of the IIIG5 epitope within the RCγ chain. 210

Figure II-4. Alignment of integrin α2A domains from different species. 212

Figure II-5. Alignment of A-domain of different human integrin α-chains. 214

Figure III-1. CD experiments. 216

Figure III-2. S-SAD electron density. 217

Figure III-3. Thermodynamic Paths for Free Energy Calculations. 218

List of Abbreviations

A2b	Adenosine A2b receptor
ADP	Adenosine Diphosphate
AHC	Anti Human Fc Capture

α2A	The A domain of Integrin α2
amp	ampicillin
ATCC	American Type Culture Collection
ATP	Adenosine Triphosphate
AUC	Analytical Ultracentrifugation
BM	Basement Membrane
BS^3	Bis(Sulfosuccinimidyl)suberate
BSA	Bovine Serum Albumin
CAM	Chorioallantoic Membrane
CCP4	Collaborative Computational Project No. 4
CD	Circular Dichroism
CD31	Cluster of Differentiation 31
CID	Collision Induced Dissociation
CL	Cross-Linked
CLEC-2	C-type Lectin-like Receptor 2

CLRP	C-type Lectin-Related Protein
Col-IV	Collagen IV
COMPcc	Cartilage Oligomerization Matrix Protein coiled-coil domain
CRD	Carbohydrate Recognizing Domain
Cryo-EM	Cryogenic Electron Microscopy
DAPI	4,6-Diamidino-2-phenylindole
DCC	Deleted in Colorectal Cancer
DFT	Density Functional Theory
DLS	Dynamic Light Scattering
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
DSCAM	Down Syndrome Cell Adhesion Molecule
E(10.5/15)	Embryonic day 10.5/15
EC	Endothelial Cell
ECGM2	Endothelial Cell Growth Medium 2

ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
EGTA	Ethylene Glycol-bis(β-aminoethyl ether)-N,N,N',N'-Tetraacetic Acid
eLBOW	electronic Ligand Builder and Optimization Workbench
ELISA	Enzyme-Linked Immunosorbent Assay
ETC	Electron Transport Chain
FBHW	Flat Bottomed Harmonic Well
Fc	Fragment crystallizable region
FCS	Fetal Calf Serum
FITC	Fluorescein Isothiocyanate
FPLC	Fast Protein Liquid Chromatography
GPI	Glycophosphatidylinositol
GPIb	Glycoprotein 1b
GPIV	Glycoprotein IV
H&E	Haemotoxylin and Eosin

HBS	HEPES Buffered Saline
HDLEC	Human Dermal Lymphatic Endothelial Cells
HDMEC	Human Dermal Microvascular Endothelial Cells
HEK293	Human Embryonic Kidney-293
HEPES	2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethanesulfonic Acid
HPLC	High Pressure Liquid Chromatography
HUVEC	Human Umbilical Vein Endothelial Cell
IMS	Ion Mobility Separation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KO	Knockout
KRMM	Kernel Ridge Mixed Model
Laser	light amplification by stimulated emission of radiation
LB	Lysogeny Broth
LE	Laminin EGF-like
LED	Light Emitting Diode

LG	Laminin Globular
Lmα1	Laminin α1
Lmβ1	Laminin β1
Lmγ1	Laminin γ1
LN	Laminin N-terminal
MBP	Maltose Binding Protein
MCTI	Multi-Configurational Thermodynamic Integration
MD	Molecular Dynamics
MS	Mass Spectrometry
MST	Microscale Thermophoresis
NCS	Non-Crystallographic Symmetry
NDSB-256	Nondetergent Sulfobetaine-256
Net1	Netrin-1
Net4	Netrin-4
Net4ΔC	Netrin-4 lacking the C terminal NTR domain

Net4-FL	Full Length Netrin-4
NMR	Nuclear Magnetic Resonance
NPS-Cl	2-Nitro-Phenylensulfenylchloride
NTR	Netrin (when referring to the netrin C-terminal domain)
OB	Olfactory Bulb
OD	Optical Density
PBS	Phosphate Buffered Saline
PBST	PBS with 0.1% Tween 20
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDFT/MM	Perturbation DFT with classical Molecular Mechanics methods
PDGF	Platelet-Derived Growth Factor
PEG	Polyethylene Glycol
PFA	Paraformaldehyde
PGDF-BB	PGDF subunit b homodimer

PMSF	Phenylmethylsulfonyl Fluoride
pNpp	para-Nitrophenyl Phosphate
PVC	Perivascular Cell
RC	Rhodocetin
RCαβ	Rhodocetin dimer of the α and β subunits
RCαβγδ	Rhodocetin tetramer of the α, β, γ, and δ subunits
RCγδ	Rhodocetin dimer of the γ and δ subunits
Rh	Hydrodynamic Radius
RHCC	Right Handed Coiled-Coil
RHCC-NT	RHCC-Nanotube
RMSD	Root Mean Square Deviation
ROI	Region of Interest
SAXS	Small Angle X-ray Scattering
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

SEC	Size Exclusion Chromatography
S-layer	Surface Layer
SOC	Super Optimal broth with Catabolite repression
SPC	Single Point Charge
S-SAD	Sulfur-Single wavelength Anomalous Dispersion
STABLE	Stalk-Associated Archaebacterial Endo-(protease)
TBS	Tris Buffered Saline
TEA	Terminal Electron Acceptor
TFA	Trifluoroacetic acid
THP	Tris(hydroxymethyl)phosphine
TPCK	Tosyl Phenylalanyl Chloromethyl Ketone
Tris	Tris(hydroxymethyl)aminomethane
Unc5	Uncoordinated 5
vBM	vascular Basement Membrane
VEGF	Vascular EGF

vWF	von Willebrand Factor
XDS	X-ray Detector Software
XRD	X-Ray Diffraction

Chapter 1: Non-enzymatic Extracellular Proteins and Their Role in Growth, Defence, and Metabolism

Introductory Remarks

The ability to interact with the environment and respond to stimuli is a hallmark of living things. These responses are often effected by secreted molecules, such as proteins; in this thesis I will present my work on three such systems where the knowledge of the structure of these proteins, either alone or in complex with other molecules, allows us to better understand their functionality. This introductory chapter will serve to provide an overview of the three systems in question; more details will be presented in their respective chapters. The first system I will discuss is the secreted guidance protein Netrin-4 (Net4), which plays a role in remodeling the basement membrane (BM) of tissues, angiogenesis, and BM permeability[1,2,3]. In this thesis I will present the high resolution structure of Net4 in combination with lower resolution complex structures, binding assays, and mutagenesis, which allows us to describe the role and mechanism of action for this protein where there was previously none. I will also discuss Rhodocetin (RC), a snake venom toxin from the Malayan pit viper[4], in complex with the I domain of Integrin α-2, essential for the blood clotting response[5], where the gross conformational changes between free and bound RC are not only academically interesting, but may have important practical applications. Finally, Right Handed Coiled-Coil (RHCC), a portion of the surface layer (S-layer) protein tetrabrachion of the archaea *Staphylothermus marinus*[6], where the structure of the protein in complex with elemental sulfur has shed new light on the metabolism of the archaea.

An Overview of Netrins and the Netrin Family of Proteins

Laminins and the Basement Membrane

Compartmentalization is a key aspect of all biology; on the most basic level, membranes separate organelles from the rest of the cytoplasm and each higher level of organization comes with a corresponding increase in exclusion and boundaries. The separation of the epithelium from the underlying tissues in higher level organisms is accomplished by the basement membrane (BM), a specialized extracellular matrix (ECM) present under the skin and around the respiratory tract, gastrointestinal tract, kidney, etc[7]. The two major structural components of the BM are collagen IV and laminin; nidogens, agrins, and perlecans are less abundant, but are also required for structural integrity. Structural similarity between laminins and collagens exists only at the most superficial level; both are trimeric proteins that crosslink to from larger superstructures which provide mechanical support to the BM[8]. Collagen IV is composed of subunits α1 through α6, with the most common variant in BMs being α1α1α2. All collagen subunits have an extended helical structure permitted by the incorporation of a glycine in every third position and a high proportion of proline and hydroxyproline[7]. These subunits assemble into the collagen triple helix, which self-associate to form longer fibrous structures that lend structural stability to much of the basement membrane. Laminins are composed of an α, β, and γ subunit that associate through a coiled-coil domain near the C-termini of the proteins. Though there exists five α, four β, and three γ variants, only 16 mammalian isoforms have been isolated to date[9]. Laminins adopt a cruciform arrangement, where the coiled-coil domain forms the base and the N-terminal domains of the subunits constitute the short arms of the cross. Polymerizing laminins, such as laminin-111 or laminin-511, can associate via the Laminin N-terminal (LN) domain. A triple complex is formed through the LN domains of an α, β, and γ subunit, each from a separate laminin trimer. This results in a 2-dimensional network of laminin that associates with the cellular membrane on the C-terminal end[8]. Non-polymerizing laminins such as laminin-411 lack the α-chain short arm, and resemble more closely a Y than a cross[10].

Laminin Domain Topology

With the exception of a few variants, laminins maintain a conserved domain topology. Exceptions include α3A, α4, and γ2, which exhibit a truncated N-terminus. The N-terminal domain, aptly named the LN or laminin N-terminal domain, is responsible for laminin polymerization. It is a 7-stranded β-sandwich jellyroll domain of approximately 250 residues. Mutation of this domain results in a number of dystrophic genetic conditions, as the incorrectly formed BM is unable to function correctly. Following the LN domain are several repeated LE or laminin-EGF-like domains. These are ~60 residue domains formed of irregular coils held together by 4 disulfide bridges. These rigid domains result in the extension of the short arms away from the coiled-coil domain and serve as binding sites for other BM proteins, such as nidogen[11]. Laminin short arms also contain the laminin B domain of unknown function. Following the multi-domain short arm is the ~500 residue coil region, which mediates the association of the laminin heterotrimer. α-laminins have additional C-terminal laminin globular (LG) domains that interact with sulfatides in the cell membrane, as well as α-dystroglycan, integrins, and other basement membrane proteins[8–10].

A Brief Overview of the Netrins

The short arm domain topology of the laminins is reflected in another class of proteins, the netrins. A class of molecules involved in regulating tissue growth, netrins are, with exceptions, composed of an LN domain, followed by 3- 3.5 LE domains, and terminated by a netrin (NTR) domain[12]. In humans, the netrin family consists of the soluble netrins 1, 3, 4, 5, and the GPI anchored G1 and G2[13]. The most thoroughly studied member of this family is netrin-1 (Net1). Net1 plays an important role in many cellular pathways; it has been implicated in cell survival, angiogenesis, tumour growth, and axon guidance[14]. The binding of Net1 to its dependence receptors can cause axon repulsion (Unc5), or attraction (DCC, neogenin)[15–17], and its role in a number of cancers, as well as neurodegenerative disorders such as multiple sclerosis and alzheimers, is the focus of many research groups. Recent data suggests that Net1 interacts with its dependence receptors through its LE domains, specifically the LE-2 domain[14,18,19]. Netrin-3 is expressed in the peripheral nervous system and is thought to regulate its development, while netrin-5 resides in the neurogenic regions of the adult brain. However, the modes of action of these two proteins is unclear. Interestingly, netrin-5 lacks the LN domain found in all other netrins[13]. The membrane anchored netrins G1 and G2 lack the NTR domain, as their interaction with the cell membrane is much more immediate. They are localized at synaptic interfaces, and are postulated to modulate synaptic activity through interaction with the aptly named netrin G ligands 1 and 2 through their LN domain[20]. Interestingly, netrin G1 knockout in mice results in the loss of either the anxiety or fear response, depending on whether the knockout is localized in cortical excitatory neurons (former) or inhibitory neurons (latter)[21].

Netrin 4

As all of the other members of the netrin family are closely associated with nervous tissue function, it is unsurprising that the relatively smaller body of literature of Net4 contains a number of papers examining a potential interaction between it and the canonical Net1 dependence receptors. However, it has been found to regulate tissue vascularization and has abnormal expression levels in a number of cancers, such as gastrointestinal and breast[1,3,22–24]. Net4 was initially designated β-netrin, due to its homology with β-laminin[12]. In contrast, all other human netrins bear a closer resemblance to γ-laminin. Our work in *Structural decoding of netrin-4 reveals unique non-enzymatic disruptive forces towards mature basement membranes* provides evidence that the regulatory function of Net4 is not due to interactions with DCC or Unc5, but rather to a novel pathway involving direct interaction with the basement membrane[25].

Haemostasis and its Antagonism by *Calloselasma rhodostoma*

Mammalian Haemostasis

For larger organisms such as mammals, direct diffusion from the environment is insufficient for the metabolic needs of the cell; we rely on a closed loop circulatory system to carry metabolites to and catabolites away from cells and tissues of the body. As the circulatory system is essential for our survival, we have evolved a complex regulatory system to ensure holes in the vasculature are plugged as efficiently as possible, but the blood in our bodies remains fluid. A central part of the blood clot response, or haemostasis, is the platelet. Much smaller than red or white blood cells, the average platelet count in humans is between 150,000 and 400,000 per microliter of blood. Platelets interact both directly and indirectly with damaged areas and, upon activation, undergo cytoskeletal rearrangement and adhere to neighbouring platelets[5,26]. Accompanying this change in morphology is the release of pro-clotting compounds such as ADP from the platelet, which activates more platelets, resulting in a cascade of activation and the formation of a blood clot[27]. Further signalling pathways result in the cleavage of fibrinogen by thrombin to form fibrin, which results in a more structurally sound mature blood clot.

The interaction of platelets at the site of injury follows two main pathways. When endothelial cells are damaged, they release a number of pro-clotting chemical cues. Among these is the ~270 kDa glycoprotein von Willebrand Factor (vWF), which is associates into multimers upwards of 250 MDa in size. The importance of vWF is evidenced by numerous genetic blood disorders centred on vWD mutations, many of which have a high mortality rate if left untreated[28]. One of the primary functions of vWF is to interact with collagen exposed by tissue damage via the vWF type A domain. Once collagen has bound, vWF can activate platelets through platelet glycoprotein Ib (GPIb)[5]. Platelets can also directly bind to exposed collagen through the integrin α2β1. Integrin α2 also contains a vWF type A domain that undergoes a conformational change upon collagen binding that ultimately results in platelet activation[29].

Venom of the Malayan Pit Viper

The Malayan Pit Viper (*Calloselasma rhodostoma*) is a venomous snake native to Southeast Asia. Though untreated bites are rarely fatal, they often result in permanent damage or even loss of function in the affected limb[4]. Envenomation may also cause a prolonged coagulation defect, which can not be treated by antivenom[30]. The snake venom itself is a cocktail of proteins and glycoproteins designed to damage the tissue of the bite victim, and to neutralize the blood clotting response, resulting in severe hemorrhage. It is composed of several classes of molecules. Rhodostoxin, a hemorrhagin, acts by destroying collagenous basement membranes and connective tissues, thereby weakening blood vessel walls and promoting hemorrhage. L-amino acid oxidase, 30% of the venom dry weight, catalyzes the formation of ammonia, hydrogen peroxide and an α-keto acid from an amino acid and water. The primary cause of tissue damage is thought to be the highly reactive oxygen, which damages membranes, proteins, and nucleic acids. Thrombin-like enzymes such as ancrod comprise 7.5% of the venom dry weight; these enzymes activate fibrinogen outside of the normal haemostatic cascade. The fibrin formed through ancrod activation forms microclots that are quickly dissolved by the native enzyme plasmin. This results in the depletion of fibrinogen in the vicinity of the bite, thereby inhibiting normal clot formation. A platelet aggregation inducer, rhodocytin, interacts with CLEC-2 on the platelet surface, inducing platelet activation outside of the normal blood clotting cascade and results in a depletion of platelets at sites of tissue damage. Fibrinolytic factors degrade fibrin and inhibit the maturation of blood clots[4,31]. The final haemostatic inhibitor of the *C. rhodostoma* venom is RC. A heterotetramer, the bifunctional RC separates into αβ and γδ dimers upon recognition of their targets on the platelet surface[32–34]. The αβ heterodimer recognizes GPIb and blocks vWF mediated platelet aggregation while the γδ heterodimers recognize the A domain of α2β1 integrin and blocks vWF independent aggregation. In *Dramatic and concerted conformational changes enable rhodocetin to block α2β1 integrin selectively,* we compare the previously crystallized RC tetramer with the complex of RCγδ and the A domain of integrin α2 and explore the conformational changes required to release the αβ subunits and bind integrin[35].

The Surface Layer and Metabolism of *Staphylothermus marinus*

Archaeal Surface Layers

Unlike bacteria, archaea lack a peptidoglycan layer and require other means to protect the integrity of the cell; instead, this is accomplished with a membrane associated glycoprotein surface layer (S-layer) which assembles to form a proteinaceous shell on the exterior of the organism[36]. The composition of the S-layer varies widely in archaea; predicted molecular masses of the individual subunits range from as little as 26kDa to over 250kDa, and have a varying number of glycosylation sites[37]. S-layers are membrane anchored, and self-assemble to form monocrystalline layers with varying symmetry. This self-assembly process has been noted in a number of isolated S-layer proteins, and has potential in applications such as molecular sieving and immobilization of functional molecules in affinity membranes, biosensors, and immunoassays[38]. Though the assembly and structure of these proteins has been examined thoroughly, there is a dearth of knowledge on the functions of these structures besides thermal, mechanical, and osmotic stabilization. S-layers are not unique to archaea; in bacteria, S-layer proteins have been implicated in virulence, avoidance of the immune system, and even to selective biomineralization[39], but parallel data is lacking for archaea.

Surface Layer of *Staphylothermus marinus*

One of the most interesting S-layer systems is that of *Staphylothermus marinus*; a hyperthermophilic sulfur-reducing archaea isolated both at deep sea hydrothermal vents (>2000m depth) and at shallow water marine hot springs (10m). The optimal growth conditions of *S. marinus* is around 94°C and a pH of 6.5 respectively. *S. marinus* is a member of the *Desulfurococcaceae* family, and grow in small, ‘grape-like’ clusters[40]. It uses peptides as carbon and energy sources and remains metabolically active even in the absence of sulfur, though it requires sulfur for growth[41]. The S-layer of *S. marinus* is composed of 3 subunits, the light chain (85 kDa), the heavy chain (92 kDa), and a unique enzymatic component, the protease STABLE (150 kDa). The heavy chain is predicted to be mostly α-helical in nature, and it is membrane anchored at the C terminus with 4 copies coming together to form a parallel coiled-coil stalk.

The N-terminus of the heavy chain contains a hinge region where the 4 chains diverge before interacting with the β-stranded light chain. This complex is known as the tetrabrachion, and it associates with adjacent tetrabrachia in a carbohydrate-dependant manner via the light chain to form a network that coats the surface of the cell. The coiled-coil domain is roughly 70nm in length and appears almost completely straight in electron micrograph images. This, in combination with the networked canopy of light chains, gives rise to a pseudo-periplasmic space[42,43]. Roughly halfway up the α-helical stalk resides 2 copies of the subtilisin-like protease, STABLE. Thus far, STABLE is the only identified stoichiometric enzymatic component in an S-layer; it is suspected that STABLE provides peptides for metabolism[44]. Though the interaction between tetrabrachion subunits is relatively easily broken, the tetrabrachion itself is incredibly difficult to disassemble, they resist denaturation even while boiling in 6M guanidinium hydrochloride. The tetrabrachion is also the source of the first discovered naturally occurring right handed coiled-coil. This motif appears about halfway up the stalk, possibly forming a portion of the STABLE binding site[6,43,45].

Elemental Sulfur Metabolism in Desulfurococcales

A common theme in extremophilic organisms, particularly hyperthermophiles, is the absence of oxygen as the terminal electron acceptor in the electron transport chain (ETC). Though the half reaction of oxygen has a very large reductive potential, increasing the ATP yield of the ETC, the solubility and therefore availability of oxygen in aqueous solutions decreases with increasing temperature, necessitating the use of alternative, non-gaseous, electron acceptors[46,47]. The order

Desulfurococcales span a number of diverse organisms, both autotrophs and heterotrophs, that can use nitrate, nitrate, thiosulfate, hydrogen, or even oxygen as terminal election acceptors[48]. *S. marinus*, like several other members of this order, is a heterotrophic sulfur reducer, which uses organic acids, such as peptides, as a carbon source. Elemental sulfur (S_8 is the most common allotrope) is used in the place of oxygen, and hydrogen sulfide is produced[49,50]. Though sulfur respiration is present in many extremophiles, little is known about the proteins that sequester sulfur from the environment or catalyze the reaction. Currently, only a single structure has been deposited in the protein database with elemental sulfur bound, that of a sulfur quinolone reductase (PDB 3HYW). Although the resolution of this structure is very good (2.0Å), the sulfur ring is poorly defined[51]. There is also currently no structural data available for the proteins that carry sulfur into the cell, and in many obligate sulfur reducers, this protein has not even been identified. In *The Right Handed Coiled-Coil nanotube in complex with elemental cyclo-octasulfur (S_8)*, the structure of RHCC with S_8 along with computational support provides perhaps the first atomic resolution structural information of how a sulfur reducer sequesters elemental sulfur from its environment[52].

References

1. Liu, Y. *et al.* Novel Role for Netrins in Regulating Epithelial Behavior during Lung Branching Morphogenesis. *Curr. Biol.* **14,** 897–905 (2004).

2. Schneiders, F. I. *et al.* Binding of netrin-4 to laminin short arms regulates basement membrane assembly. *J. Biol. Chem.* **282,** 23750–23758 (2007).

3. Nacht, M. *et al.* Netrin-4 regulates angiogenic responses and tumor cell growth. *Exp. Cell Res.* **315,** 784–794 (2009).

4. Tang, E. L. H., Tan, C. H., Fung, S. Y. & Tan, N. H. Venomics of Calloselasma rhodostoma, the Malayan pit viper: A complex toxin arsenal unraveled. *J. Proteomics* **148,** 44–56 (2016).

5. Nuyttens, B. P., Thijs, T., Deckmyn, H. & Broos, K. Platelet adhesion to collagen. *Thromb. Res.* **127,** S26–S29 (2011).

6. Stetefeld, J. *et al.* Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat. Struct. Biol.* **7,** 772–6 (2000).

7. LeBleu, V. S., MacDonald, B. & Kalluri, R. Structure and Function of Basement Membranes. *Exp. Biol. Med.* **232,** 1121–1129 (2007).

8. Sasaki, T., Fässler, R. & Hohenester, E. Laminin: The crux of basement membrane assembly. *J. Cell Biol.* **164,** 959–963 (2004).

9. Schéele, S. *et al.* Laminin isoforms in development and disease. *J. Mol. Med.* **85,** 825–836 (2007).

10. Hohenester, E. & Yurchenco, P. D. Laminins in basement membrane assembly. *Cell Adh. Migr.* **7,** 56–63 (2013).

11. Purvis, A. & Hohenester, E. Laminin network formation studied by reconstitution of ternary nodes in solution. *J. Biol. Chem.* **287,** 44270–44277 (2012).

12. Koch, M. *et al.* A novel member of the netrin family, beta-netrin, shares homology with the beta chain of laminin: identification, expression, and functional characterization. *J. Cell Biol.* **151,** 221–234 (2000).

13. Moore, S. W., Tessier-Lavigne, M. & Kennedy, T. E. Netrins and Their receptors. in *Axon Growth and Guidance* (ed. Bagnard, D.) 17–31 (Springer New York, 2007). doi:10.1007/978-0-387-76715-4_2

14. Grandin, M. *et al.* Structural Decoding of the Netrin-1/UNC5 Interaction and its Therapeutical Implications in Cancers. *Cancer Cell* **29,** 173–185 (2016).

15. Dun, X. P. & Parkinson, D. B. Role of Netrin-1 signaling in nerve regeneration. *Int. J. Mol. Sci.* **18,** 1–22 (2017).

16. Finci, L., Zhang, Y., Meijers, R. & Wang, J. H. Signaling mechanism of the netrin-1 receptor DCC in axon guidance. *Prog. Biophys. Mol. Biol.* **118,** 153–160 (2015).

17. Laumonnerie, C., Da Silva, R. V., Kania, A. & Wilson, S. I. Netrin 1 and Dcc signalling are required for confinement of central axons within the central nervous system. *Development* **141,** 594–603 (2014).

18. Finci, L. I. *et al.* The Crystal Structure of Netrin-1 in Complex with DCC Reveals the Bifunctionality of Netrin-1 As a Guidance Cue. *Neuron* **83,** 839–849 (2014).

19. Xu, K. *et al.* Structures of netrin-1 bound to two receptors provide insight into its axon guidance mechanism. *Science (80-.).* **344,** 1275–1279 (2014).

20. Brasch, J., Harrison, O. J., Ahlsen, G., Liu, Q. & Shapiro, L. Crystal structure of the ligand binding domain of netrin G2. *J. Mol. Biol.* **414,** 723–734 (2011).

21. Zhang, Q. *et al.* Netrin-G1 regulates fear-like and anxiety-like behaviors in dissociable neural circuits. *Sci. Rep.* **6,** 28750 (2016).

22. Yuan, Y. *et al.* Netrin-4 is upregulated in breast carcinoma effusions compared to corresponding solid tumors. *Diagn. Cytopathol.* **39,** 562–566 (2011).

23. Lambert, E., Coissieux, M. M., Laudet, V. & Mehlen, P. Netrin-4 acts as a pro-angiogenic factor during zebrafish development. *J. Biol. Chem.* **287,** 3987–3999 (2012).

24. Lejmi, E. *et al.* Netrin-4 inhibits angiogenesis via binding to neogenin and recruitment of Unc5B. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 12491–6 (2008).

25. Reuten, R. *et al.* Structural decoding of netrin-4 reveals a regulatory function towards mature basement membranes. *Nat. Commun.* **7,** 13515 (2016).

26. Yip, J., Shen, Y., Berndt, M. C. & Andrews, R. K. Primary platelet adhesion receptors. *IUBMB Life* **57,** 103–108 (2005).

27. Daniel, J. L. *et al.* Molecular Basis for ADP-induced Platelet Activation. *J. Biol. Chem.* **273,** 2024–2029 (1998).

28. Leebeek, F. W. G. & Eikenboom, J. C. J. Von Willebrand's Disease. *N. Engl. J. Med.* **375,** 2067–2080 (2016).

29. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J. & Liddington, R. C. Structural Basis of Collagen Recognition by Integrin α2β1. *Cell* **101,** 47–56 (2000).

30. Kini, R. M., Clemetson, K. J., Markland, F. S., McLane, M. A. & Morita, T. Toxins and hemostasis: From bench to bedside. in *Toxins and Hemostasis: From Bench to Bedside* 1–797 (2011). doi:10.1007/978-90-481-9295-3

31. Kong, C. & Chung, M. C. M. Purification and characterization of a variant of rhodocetin from Calloselasma rhodostoma (malayan pit viper) venom. *J. Protein Chem.* **20,** 383–390 (2001).

32. Eble, J. a *et al.* The alpha2beta1 integrin-specific antagonist rhodocetin is a cruciform, heterotetrameric molecule. *FASEB J.* **23,** 2917–2927 (2009).

33. Eble, J. a & Tuckwell, D. S. The alpha2beta1 integrin inhibitor rhodocetin binds to the A-domain of the integrin alpha2 subunit proximal to the collagen-binding site. *Biochem. J.* **376,** 77–85 (2003).

34. Navdaev, A., Lochnit, G. & Eble, J. A. The rhodocetin αβ subunit targets GPIb and inhibits von Willebrand factor induced platelet activation. *Toxicon* **57,** 1041–1048 (2011).

35. Eble, J. A. *et al.* Dramatic and concerted conformational changes enable rhodocetin to block α2β1 integrin selectively. *PLoS Biol.* **15,** e2001492 (2017).

36. Baumeister, W. & Lembcke, G. Structural features of archaebacterial cell envelopes. *J. Bioenerg. Biomembr.* **24,** 567–575 (1992).

37. Engelhardt, H. & Peters, J. Structural Research on Surface Layers: A Focus on Stability, Surface Layer Homology Domains, and Surface Layer–Cell Wall Interactions. *J. Struct. Biol.* **124,** 276–302 (1998).

38. Sleytr, U. B. & Sára, M. Bacterial and archaeal S-layer proteins: Structure-function relationships and their biotechnological applications. *Trends Biotechnol.* **15,** 20–26 (1997).

39. Sleytr, U. B. & Beveridge, T. J. Bacterial S-layers. *Trends Microbiol.* **7,** 253–260 (1999).

40. Fiala, G., Stetter, K. O., Jannasch, H. W., Langworthy, T. A. & Madon, J. Staphylothermus marinus sp. nov. Represents a Novel Genus of Extremely Thermophilic Submarine Heterotrophic Archaebacteria Growing up to 98 °C. *Syst. Appl. Microbiol.* **8,** 106–113 (1986).

41. Hao, X. & Ma, K. Minimal sulfur requirement for growth and sulfur-dependent metabolism of the hyperthermophilic archaeon Staphylothermus marinus. *Archaea* **1,** 191–7 (2003).

42. Peters, J. *et al.* Tetrabrachion: A Filamentous Archaebacterial Surface Protein Assembly of Unusual Structure and Extreme Stability. *J. Mol. Biol.* **245,** 385–401 (1995).

43. Peters, J., Baumeister, W. & Lupas, A. Hyperthermostable Surface Layer Protein Tetrabrachion from the ArchaebacteriumStaphylothermus marinus: Evidence for the Presence of a Right-handed Coiled Coil Derived from the Primary Structure. *J. Mol. Biol.* **257,** 1031–1041 (1996).

44. Mayr, J. *et al.* A hyperthermostable protease of the subtilisin family bound to the surface layer of the archaeon Staphylothermus marinus. *Curr. Biol.* **6,** 739–49 (1996).

45. Lupas, A. Predicting coiled-coil regions in proteins. *Curr. Opin. Struct. Biol.* **7,** 388–393 (1997).

46. Barton, L. L., Fardeau, M.-L. & Fauque, G. D. Hydrogen Sulfide: A Toxic Gas Produced by Dissimilatory Sulfate and Sulfur Reduction and Consumed by Microbial Oxidation. in 237–277 (2014). doi:10.1007/978-94-017-9269-1_10

47. Schut, G. J., Boyd, E. S., Peters, J. W. & Adams, M. W. W. The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. *FEMS Microbiol. Rev.* **37,** 182–203 (2013).

48. Gupta, R. S. & Shami, A. Molecular signatures for the Crenarchaeota and the Thaumarchaeota. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **99,** 133–157 (2011).

49. Liu, Y., Beer, L. L. & Whitman, W. B. Sulfur metabolism in archaea reveals novel processes. *Environ. Microbiol.* **14,** 2632–2644 (2012).

50. Boyd, E. S. *et al.* Isolation, characterization, and ecology of sulfur-respiring Crenarchaea inhabiting acid-sulfate-chloride-containing geothermal springs in Yellowstone National Park. *Appl. Environ. Microbiol.* **73,** 6669–6677 (2007).

51. Marcia, M., Ermler, U., Peng, G. & Michel, H. The structure of Aquifex aeolicus sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. *Proc. Natl. Acad. Sci.* **106,** 9625–9630 (2009).

52. McDougall, M. *et al.* Archaea S-layer nanotube from a ‘black smoker’ in complex with cyclo-octasulfur S 8 rings. *Proteins Struct. Funct. Bioinforma.* (2017). doi:10.1002/prot.25385

Chapter 2: Instrumentation and Methods

Preparation of RHCC-6His *E. coli*

Due to low expression levels and difficulty of culturing native organisms, many proteins are recombinantly expressed. In the case of RHCC, *Escherichia coli* (*E. coli*) strain BL21 from New England Biolabs was used as an expression system due to its rapid growth rate and well characterized nutrient requirements. A previously designed RHCC construct (kindly gifted by Suat Özbek of the University of Heidelberg, Germany) containing a thrombin cleavable N-terminal 6His tag in an IPTG inducible ampicillin resistant pET15b vector was transformed into *E. coli* by mixing ~5ng of plasmid with ~10^{11} cells in 50 µL 15% glycerol and subjecting the mixture to 1800V in a 0.1cm gap electroporation cell. Cells were then allowed to recover in SOC media for 1 hour at 37°C prior to plating onto LB agar containing 100 µg/mL ampicillin. The plates were incubated at 37°C overnight. 50 mL of LB broth was inoculated with a single colony picked from the plate and allowed to grow at 37°C for 12 hours with aeration. The resulting cells were collected and centrifuged at 3000x*g* to pellet the cells which were then resuspended in 50% glycerol and frozen at -70°C. These cells form the stock for all subsequent RHCC expression experiments.

Expression and Purification of RHCC

A pre-culture of RHCC-6His *E. coli* inoculated into 50mL of LB broth with 100 μg/mL ampicillin (LB +amp) was grown overnight at 37°C with aeration; this culture was then added to 1.6L of LB +amp and grown for approximately 2 hours until the optical density at 600nm reached 0.6. IPTG is added to a final concentration of 0.5mM and the cells are allowed to grow for a further 3 hours. Cells were then harvested and pelleted at 12 000x*g* and frozen at -70°C to begin the lysis process. Cells were then thawed on ice and resuspended in a buffer containing 20mM phosphate, 2mM imidazole pH=7, 6M guanidine hydrochloride. Cell lysis was furthered on ice in a dounce style homogenizer and completed by 2 minutes of sonication at 50% amplitude in a Fisher Scientific Model 500 Sonic Dismembrator. The resulting lysate was then centrifuged at 70 000x*g* at 6°C to remove insoluble cellular components. The 6His tag on the RHCC construct has a high affinity to bivalent colbalt or nickel ions[1], and was used to separate the tagged RHCC from the other cellular components. The soluble lysate was passed through a Talon cobalt column (GE healthcare) and washed first with the denaturing resuspension buffer, then with a buffer containing 20mM phosphate, 2mM imidazole pH=7 to allow the protein to refold on the column. Purified RHCC-6His was then eluted from the column with 150mM imidazole, pH=7, 300mM NaCl. All column purification was done at 4°C. RHCC-6His is then dialyzed into a buffer containing 20mM Tris, pH= 8.4, I=154mM (NaCl) before cleavage with bovine thrombin (GE healthcare). Thrombin was removed from the solution by passage through a benzamidine serine protease affinity column (GE healthcare). Purity and homogeneity of the thrombin digested protein was assayed by SDS-PAGE and DLS and the pure RHCC was stored at 4°C until required.

Dynamic Light Scattering

Dynamic Light Scattering (DLS) is a non-destructive method used to measure the hydrodynamic radius (r_H) of a particle in solution, and can be used to determine solution monodispersity or monitor particle size as a function of temperature or concentration[2]. We performed DLS on Net4 at concentrations between 0.5 and 6 mg/mL, and on RHCC at different temperatures to determine if oligomerization changed with these variables, or if aggregation occurs. Solution monodispersity is key for a number of structural biology techniques. Data was collected in a 12 μL quartz cuvette using a Zetasizer Nano S system (Malvern Instruments), which utilizes a 4mW 633nm laser and a scattering angle of 173°. The Zetasizer software calculates the diffusion constant using an autocorrelation function, which is then used to calculate a hydrodynamic radius using the Stokes-Einstein relationship. It is important to note that the hydrodynamic radius does not necessarily match any physical dimension of the scattering particle, it is merely the radius of an ideal sphere that moves with the same diffusion constant of the particle in solution[3].

Circular Dichroism

Circular Dichroism Spectroscopy (CD) is a technique used to determine the secondary structure of a protein. Chiral chromophores will differentially absorb left- and right- handed circularly polarized light. Aside from the intrinsic chirality of the individual amino acids, higher order protein folds such as α-helices and β-strands have degrees of chirality that differ from an unfolded or random coil structure. As peptide bonds absorb strongly at ~200nm, the absorbance difference between 180 and 260nm has been determined for many secondary structure elements, and so it is possible to deconvolute a CD spectrum and determine the contribution from each secondary structural element to the overall fold[4]. Though RHCC is known to be entirely α-helical from previous unpublished CD experiments at 20°C, it was important to ascertain if RHCC maintains its folded state at the elevated temperatures used for the sulfur incorporation. Therefore, CD spectra were collected in triplicate on a J-810 spectropolarimeter (Jasco) at 20 and 81°C on 0.2mg/mL RHCC in 12mM phosphate, pH 7.0, 200 mM NaF in a 0.1 cm quartz cuvette (Hellma) with a scan speed of 7.5nm/min.

X-ray Crystallography

Introduction to Macromolecular Crystallography

The function of biological macromolecules depends less on their chemical composition and more on the 3 dimensional structure and how that structure interacts with other molecules in a biological context. Determination of the 3D structure can be done using Nuclear-Magnetic Resonance Spectroscopy (NMR), Cryo-Electron Microscopy (Cryo EM) or X-ray diffraction (XRD); each technique has its own advantages and disadvantages[5,6]. XRD has no theoretical size limitations, does not require labels, and has many well polished software suites with very active communities behind them. The main disadvantage is that XRD requires a crystalline substrate, and it can be very difficult to induce crystallization in biological molecules. As a consequence, XRD does not work well with intrinsically disordered or extremely dynamic systems.

Preparation and Crystallization of RHCC-S_8

Initial attempts to prepare RHCC-S_8 involved solubilization of the extremely hydrophobic elemental sulfur (Fisher Scientific) using sodium dodecyl sulfate (SDS) (Anachemia), followed by incubation with 11mg/mL RHCC for 3 days. Removal of SDS by dialysis failed, as did passage through a Superdex 75 size exclusion column (GE healthcare). DLS results yielded a particle dimension too large to be attributed to RHCC alone. Instead, a new protocol was developed where ~15 mg of sublimed sulfur was added to 370 μL of 8.15 mg/mL RHCC in 10 mM bicine buffer (pH=8.0, I=154 mM (NaCl)) (Fisher Scientific) and heated to 80°C with constant stirring. The solution was held at this temperature for 4 days, followed by removal of excess sublimed sulfur by centrifugation and filtration. DLS analysis of the resulting supernatant yielded an r_H consistent with that of an RHCC tetramer. The protein was concentrated to 11mg/mL and crystals were grown by hanging drop vapour diffusion at 4°C. Specifically, 2 μL of protein and 2 μL of the reservoir solution were mixed and suspended over a 1 mL reservoir. The reservoir solution that yielded crystals was composed of 1.6M ammonium sulfate (Fisher Scientific) and 0.1M Tris pH=8.0 (Fisher Scientific). Crystals appeared after 1 week.

Crystallization of Netrin-4ΔC

Net4ΔC was crystallized by hanging drop vapour diffusion at 20°C. 2 μL of 1.2 mg/mL of Net4ΔC in 50mM Tris pH=7.5, 200mM NaCl was mixed with 1.6 μL reservoir solution (20% PEG 3350, 0.2M calcium acetate, 0.1M sodium cacodylate pH=6.5) and 0.4 μL NDSB-256 from the HR2-428 additive screen (Hampton Research). Crystals appeared after 1 week.

Data Collection, Processing and Structure Solution of RHCC-S_8

All crystal data was collected at 100K to reduce radiation damage to the protein. To prevent ice crystal formation from damaging the protein crystal and degrading the signal to noise ratio, crystals were soaked in a 4°C solution of mother liquor with 20% glycerol v/v prior to flash freezing in liquid nitrogen. Images were collected from 0-110° in 1° wedges with 10 minutes exposure on a Rigaku MM-007HF diffractometer with an X-ray wavelength of 1.54178Å using an R-axis IV++ detector. No radiation damage was observed during data collection. Data was indexed integrated, and scaled with MOSFLM[7] and the CCP4-package[8,9]. Phases were calculated with rigid body refinement in Refmac5[10,11] using a polyserine RHCC model based on a previously solved RHCC structure[12] (PDB code 1YBK). Rfree flags were transferred from the model dataset to minimize model bias. The structure was built and refined using Coot[13], Refmac5, and the Phenix software package[14]. Chemical restraints for the S_8 ring were generated using JLigand[15] and eLBOW[16].

Anomalous Sulfur Signal

To obtain experimental phase information and to confirm the presence of sulfur in the structure, anomalous data was collected at the 08B1-1 beamline at the Canadian Light Source with a wavelength of 1.77123Å. At this wavelength Friedel's Law is broken for sulfur; inverted reflections no longer have the same amplitude and the difference in amplitude is entirely from the anomalous diffracting elements[5]. Crystals were again cryoprotected with a 20% glycerol solution prior to collection of 0.5° wedges with 5.0s exposures at 100K. 5 datasets from 3 different crystals were indexed and integrated with XDS[17,18], then scaled together with the CCP4-package[9]. Anomalous difference maps were calculated using phases from a previously solved RHCC structure[19] (PDB code 1FE6) without sulfur (ϕ_{calc}) and experimental amplitudes (F_{ano}) determined by subtracting F^- from F^+.

Data Collection, Processing and Structure Solution of Net4ΔC

Net4ΔC crystals were soaked with 10% ethylene glycol in mother liquor for 10 minutes prior to flash freezing in liquid nitrogen. A full cohort of 360 1° images were collected with 9 minutes exposure each on a Rigaku MM-007HF diffractometer with an X-ray wavelength of 1.54178Å, using an R-axis IV++ detector. Data was indexed, integrated, and scaled using the HKL2000[20] software suite. Phases were determined by molecular replacement with Phaser[21] using a polyserine homology model generated by Swiss-model[22] based on the previously solved laminin β structure[23] (PDB code 4AQS). The model was built and refined using Coot[13], the Phenix software package[14], and Refmac5[10,11].

Data Collection, Processing and Structure Solution of Rhodocetin-γδ in complex with the Integrin α2A domain

The triple complex of RC-γδ subunits and the Integrin α2A domain was crystallized by hanging drop vapour diffusion with 2 μL of 10 mg/mL RCγδ-α2A and 2 μL of reservoir solution (2.65M ammonium sulfate, 0.1M Tris pH=8.0) over 1 mL reservoir. Crystals appeared after 6 weeks at 20°C and were soaked in mother liquor with 20% glycerol prior to flash freezing in liquid nitrogen. Data was collected in 0.5° wedges with 1.0s exposure for 194° at the 08ID-1 beamline at the Canadian Light Source. The wavelength used was 0.97949Å. Data was indexed, integrated, and scaled using XDS[17,18] and the CCP4-package[8,9]. Phases were obtained by molecular replacement using Phaser[21] with a polyserine model based on a previously solved rhodocetin structure[24] (PDB code 3GPR). The structure was built and refined using Coot[13] and the Phenix software package[14].

References

1. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D. Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Nat. Biotechnol.* **6,** 1321–1325 (1988).

2. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* **8,** 409–427 (2016).

3. Harding, S. E. & Jumel, K. Light Scattering. in *Current Protocols in Protein Science* 7.8.1-7.8.14 (John Wiley & Sons, Inc., 1998). doi:10.1002/0471140864.ps0708s11

4. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89,** 392–400 (2008).

5. Rupp, B. Biomolecular crystallography : principles, practice, and application to structural biology. (Garland Science, 2010).

6. Czarnocki-Cieciura, M. & Nowotny, M. Introduction to high-resolution cryo-electron microscopy. *Postepy Biochem.* **62,** 383–394 (2016).

7. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 271–281 (2011).

8. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 235–242 (2011).

9. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 72–82 (2006).

10. Vagin, A. A. *et al.* REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2184–2195 (2004).

11. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255 (1997).

12. Özbek, S., Müller, J. F., Figgemeier, E. & Stetefeld, J. Favourable mediation of crystal contacts by cocoamidopropylbetaine (CAPB). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61,** 477–480 (2005).

13. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).

14. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).

15. Lebedev, A. A. *et al.* JLigand: A graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68,** 431–440 (2012).

16. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65,** 1074–1080 (2009).

17. Kabsch, W. Xds. Acta Crystallogr. Sect. D Biol. Crystallogr. **66,** 125–132 (2010).

18. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 133–144 (2010).

19. Stetefeld, J. *et al.* Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat. Struct. Biol.* **7,** 772–6 (2000).

20. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).

21. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

22. Biasini, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42,** 252–258 (2014).

23. Carafoli, F., Hussain, S.-A. & Hohenester, E. Crystal Structures of the Network-Forming Short-Arm Tips of the Laminin β1 and γ1 Chains. *PLoS One* **7,** e42473 (2012).

24. Eble, J. a *et al.* The alpha2beta1 integrin-specific antagonist rhodocetin is a cruciform, heterotetrameric molecule. *FASEB J.* **23,** 2917–2927 (2009).

Chapter 3: Structural decoding of netrin-4 reveals unique non-enzymatic disruptive forces towards mature basement membranes

Raphael Reuten[1,2±], Trushar R. Patel[3,4±], Matthew McDougall[5±], Nicolas Rama[6], Denise Nikodemus[1,2], Benjamin Gibert[6], Jean-Guy Delcros[6], Carina Prein[7,8,9], Markus Meier[5], Stéphanie Metzger[10], Zhigang Zhou[11,12], Jennifer Kaltenberg[1,2], Karen K. McKee[13], Tobias Bald[14,15], Thomas Tüting[14,15], Paola Zigrino[16], Valentin Djonov[17], Wilhelm Bloch[18], Hauke Clausen-Schaumann[7,9], Ernst Poschl[11], Peter D. Yurchenco[13], Martin Ehrbar[10], Patrick Mehlen[6*], Jörg Stetefeld[5*], and Manuel Koch[1,2*]

[±]co-first authors

*co-senior authors

[1]Institute for Dental Research and Oral Musculoskeletal Biology, Medical Faculty, University of Cologne, Joseph-Stelzmann-Str. 52, 50931 Cologne, Germany.

[2]Center for Biochemistry, Medical Faculty, University of Cologne, Cologne, Joseph-Stelzmann-Str. 52, 50931 Cologne, Germany.

[3]Alberta RNA Research and Training Institute, Department of Chemistry and Biochemistry, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4, Canada.

[4]School of Biosciences, University of Birmingham, Edgbaston, B15 2TT, United Kingdom.

[5]Department of Chemistry, University of Manitoba, 144 Dysart Road, Winnipeg, R3T 2N2, Canada.

[6]Apoptosis, Cancer and Development Laboratory, Equipe labellisée ‘La Ligue’, LabEx DEVweCAN, Centre de Recherche en Cancérologie de Lyon, INSERM U1052-CNRS UMR5286, Université de Lyon, Centre Léon Bérard, 69008 Lyon, France.

[7]Center for Applied Tissue Engineering and Regenerative Medicine – CANTER, Munich University of Applied Sciences, Lothstrasse 34, D-80335 Munich, Germany.

[8]Laboratory of Experimental Surgery and Regenerative Medicine – ExperiMed, Department of Surgery, Clinical Center University of Munich, Nussbaumstrasse 20, D-80336, Munich, Germany.

[9]Center for Nanoscience – CeNS, Geschwister-Scholl-Platz 1, D-80539 Munich, Germany.

[10]Laboratory for Cell and Tissue Engineering, Department of Obstetrics, University Hospital Zurich, Schmelzbergstr. 12, 8091 Zurich, Switzerland.

[11]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom.

[12]Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK.

[13]Department of Pathology, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, USA.

[14]Department of Dermatology, University Hospital Magdeburg, 39120 Magdeburg, Germany.

[15]Laboratory of Experimental Dermatology, Department of Dermatology and Allergy, University of Bonn, 53105 Bonn, Germany.

[16]Department of Dermatology and Venerology, University of Cologne, 50931 Cologne, Germany.

[17]Institute of Anatomy, University of Bern, Baltzerstrasse 2, 3000 Bern, Switzerland.

[18]Institute of Cardiovascular Research and Sport Medicine, Department of Molecular and Cellular Sport Medicine, German Sport University Cologne, 50933 Cologne, Germany.

DOI 10.1038/ncomms13515

As per the Publication Ethics statement of the Nature Publishing Group, all material from this publication may be freely reproduced in an academic thesis.

Preface

In a monumental collaboration between 25 researchers from 18 institutes in Europe and North America, this article describes the structure, function, and the physiological consequences of the function of a previously under-examined extracellular matrix protein. Solution as well as crystallographic structural data provided a basis for further study into the functionally significant epitopes of Netrin-4, ultimately providing evidence that non-enzymatic basement membrane modulation produces the angiogenic and axon outgrowth effects that were previously and incorrectly attributed to the binding of Netrin-4 to canonical netrin receptors. The manuscript was accepted by *Nature Communications* on October 9, 2016.

Contributions of Authors

Matthew McDougall (University of Manitoba) collected the X-ray data and solved the structure of Netrin 4 with protein provided by Raphael Reuten (University of Cologne), who produced recombinant proteins. In collaboration with Carina Prein (Munich University of Applied Sciences, University of Munich, and Center for Nanoscience-Munich) and Hauke Clausen-Schaumann (Munich University of Applied Sciences, Center for Nanoscience-Munich) Raphael Reuten also collected AFM data of the Matrigel matrix. Tumor treatment assays were performed by Tobias Bald (University Hospital Magdeburg, University of Bonn), Thomas Tüting (University Hospital Magdeburg, University of Bonn), and Paola Zigrino (University of Cologne). Trushar R. Patel (University of Lethbridge, University of Birmingham) and Matthew McDougall collected and analyzed the SAXS data and performed AUC experiments along with Markus Meier (University of Manitoba). Denise Nikodemus (University of Cologne) performed and analyzed SPR measurements. Jean-Guy Delcros (Centre de Recherche en Cancérologie de Lyon) performed Octet binding analysis. Zhigang Zhou (University of East Anglia) and Ernst Poschl (University of East Anglia) performed the EC/PVC co-culture assay. Valentin Djonov (University of Bern) performed TEM analyses. Karen K. McKee (Robert Wood Johnson Medical School) and Peter D. Yurchenco (Robert Wood Johnson Medical School) performed the basement membrane assembly assay. Wilhelm Bloch (University of Cologne) performed the EM analysis. The manuscript was written by Matthew McDougall, Raphael Reuten, Trushar R. Patel, Markus Meier, Patrick Mehlen, Manuel Koch, and Jörg Stetefeld.

Abstract

Netrins form a family of laminin-related molecules that have been proposed to act as guidance cues either during nervous system development or the establishment of the vascular system. This guidance function has been clearly demonstrated for netrin-1 via its interaction with the transmembrane receptors DCC and UNC5 (i.e., UNC5A, B, C, and D). However, mainly due to shared homologies with netrin-1, netrin-4 was proposed to play a role in neuronal outgrowth and in developmental and pathological angiogenesis via interactions with netrin receptors. Here, we present the high-resolution structure of netrin-4 which shows unique features in comparison to netrin-1. Using a structure-function analysis, we show that netrin-4 disrupts laminin networks and basement membranes on the cell surface through high-affinity binding to the laminin γ1 chain. We hypothesize that this laminin-related function is essential for the previously described effects on axon growth promotion and angiogenesis. Our study characterizes netrin-4 as a non-enzymatic extracellular matrix protein actively disrupting pre-existing basement membranes.

Introduction

The netrin protein family is comprised of six members, the secreted netrins -1, -3, -4, -5 and the glycosylphosphatidylinositol (GPI)-anchored netrins-G1 and -G2 in mammals and exhibit homology to the N-terminal domains (LN-LEa1-3) of laminin short arms[1-8]. Secreted netrins-1,-3 and -5 are homologous to the laminin γ1 chain, whereas netrin-4 (Net4) shares a higher identity to the β1 chain (50%). Netrin-1 (Net1) has been studied in detail and the name netrin, from “the one who guides” in Sanskrit, was given to this family of laminin-related molecules because of Net1 activity[2]. Net1, one of the first studied secreted guidance factors, plays a major role in promoting both axon outgrowth and axon orientation during nervous system development[9], as well as in developmental angiogenesis. Given the structural and functional similarities between the axon growth cone and the endothelial tip cell[10] in both systems netrin-1 function is due to its ability to engage the receptors DCC and UNC5H (i.e., UNC5A, UNC5B, UNC5C and UNC5D)[11-13] and possibly other receptors such as neogenin, A2b, some integrins or DSCAM[14-17]. Due to the similarity of Net1 and Net4, the majority of studies of Net4 have been focused on describing functions of Net4 based on an assumed interaction of Net4 with DCC or UNC5 with a guidance function. Works include reported neuronal guidance function[3] and pro- as well as anti-angiogenic activity[18-20]. The most prominent *in vitro* function of Net4 is the inhibition of endothelial tube-like structure formation when Net4 is applied to endothelial cells seeded on Matrigel[20]. In addition, different cancer cells transfected with Net4 form less vascularized solid tumors[18,19]. In most cases, it was speculated that Net4 acts mainly via binding to Net1 receptors of the DCC/neogenin family and UNC5 family members[19-21]. Recently, the binding epitopes of Net1 for DCC, neogenin, and UNC5B were identified[22-24]; however, the proposed binding sites are not conserved within Net4. Of interest, even though these studies have received little attention, it was proposed that Net4 binds to the laminin γ1 chain[25,26]. Together with the fact that most *in vitro* effects shown with Net4 were obtained from culture in reconstituted matrix such as Matrigel mainly composed of laminin 111, we hypothesized that the Net4 by interacting with laminin γ1 could impact on the Matrigel *in vitro* and on the BM *in vivo*.

The BM is a specialized extracellular matrix mainly composed of laminin and collagen IV polymers[27] as well as a number of additional proteins. This matrix provides a substrate for cell attachment, serves as a barrier, and is a source of growth factors. As an example, in the vasculature the BM is located at the basal side of the endothelium. The vascular BM (vBM) assembles from proteins produced by endothelial as well as perivascular cells (PVCs)[28]. Continuous formation of the vBM along the endothelium stabilizes newly formed vessels and promotes their maturation[28,29]. Ablation of the BM core component collagen IV leads to vascular defects, hemorrhages, and embryonic lethality at E10.5 in mouse[30]. However, these mice still contain BM-like sheets assembled through laminins[30]. Although, the ultrastructural morphology of BMs is similar between different tissues, variations in protein composition are seen depending on location and during development. Laminins 411 and 511 are the predominant isoforms in the vBM formed by the β1, γ1, and either the α4 or the α5 chain. Laminin 411 is expressed in the vasculature during early developmental stages and deficiency of the laminin α4 chain leads to a vascular phenotype with transient perinatal hemorrhages[31]. In contrast, laminin 511 is found in the vBM of adult mice[32]. Laminin 511 is distinct from laminin 411 in that it contains laminin N-terminal globular (LN) domains on each of its short arms allowing its assembly into a network that is bound to endothelial cell surfaces through integrins, sulfatides, and α-dystroglycan. Following laminin assembly at cell surfaces, collagen IV is integrated into the BM[33-35].

Here, we present the high-resolution structure of Net4 showing unique features that are absent in Net1 as well as in the N-terminal laminin structures and identify the binding epitopes required for a stable Net4/laminin γ1 heterodimeric complex. Based on this high-resolution structure, we designed Net4 mutants that negatively impede interaction with laminin γ1, possibly explaining previously observed *ex vivo* effects on axon outgrowth modulation, and angiogenic activity. Our data also elucidate a mechanism by which Net4 influences the laminin network and play a vital role in angiogenesis. We also propose that the anti-angiogenic property of Net4 *in vivo* is directly related to its impact on BM rearrangement.

Results

Net4 has unique structural features absent in other netrins

In order to establish the structure-function relationship of Net4, we solved the crystal structure of Net4 lacking the C-terminal domain (Net4-ΔC) at 3.07 Å, revealing a head to stalk arrangement of 175 Å in length (**Figure 3-1a, b**). The globular shaped N-terminal domain (LN) forms the head and the three and a half rod-like consecutive LE domains make up the stalk (**Figure 3-1a, b**). The LN domain forms a β-sandwich jelly-roll motif with a front face composed of sheet S5-S2-S7 and a dorsal face composed of sheet S4-S3-S6-S1. The topology of this domain is similar to the N-terminal domain of Net1[22-24]; however, the closest structural homologue is the short arm of laminin β1 (**Appendix 1, Figure I-1a, b**). Significant structural features make this structure unique among netrins. The LN domain is stabilized by a complex pattern of disulfide bridges composed of 12 cysteine residues (**Figure 3-1a**, **b**). Two disulfide bridges between cysteine residues 72-236 and 191-234, absent in any current netrin or laminin structure, hold the S6-S7 loop in place (**Figure 3-1b**). This loop is adjacent to loop rS1-rS2, which contains the conserved Ser65 (**Appendix 1, Figure I-2a**). Interestingly, in laminin β1 (Lmβ1), the replacement of this conserved serine by an arginine allows for the formation of a binary complex between β1 and γ1 laminins; however, the ternary complex formation with the laminin α chain is abolished. Furthermore, there is clear electron density for glycan additions at three asparagine-linked glycosylation sites, Asn56, Asn163, and Asn353. The first two glycan chains are located at opposite edges of the LN domain, while Asn353 is located at the loop *b* helix of LE2 and extends towards the bottom of the LN domain (**Figure 3-1a**, **b**). Remarkably, Net4 is the first member of the netrin family with no N-linked glycan attachment at the dorsal face of the N-terminal domain. Net4-ΔC has a total accessible surface area of about 28,000 Å^2 and reveals extensive inter-domain contacts between the LN domain and subdomain LE1 (**Figure 3-1a**, **b**). Loop *b* of the LE1 domain has a large buried surface area of ~1,700 Å^2 with the lateral edge of the LN domain (**Figure 3-1c**). In addition, loop *d* of the LE1 domain is flanked by the Asn353-linked glycan attachment and trapped between the N-terminal segment and the basal helix of the LN domain (**Figure 3-1d**). Together, these extensive contact areas may limit the rotational freedom of the LN domain in comparison to other netrins. Each of the individual LE-subdomains adopts the classical LE-fold[36] consisting of irregular coil segments and forms a linear extended structure.

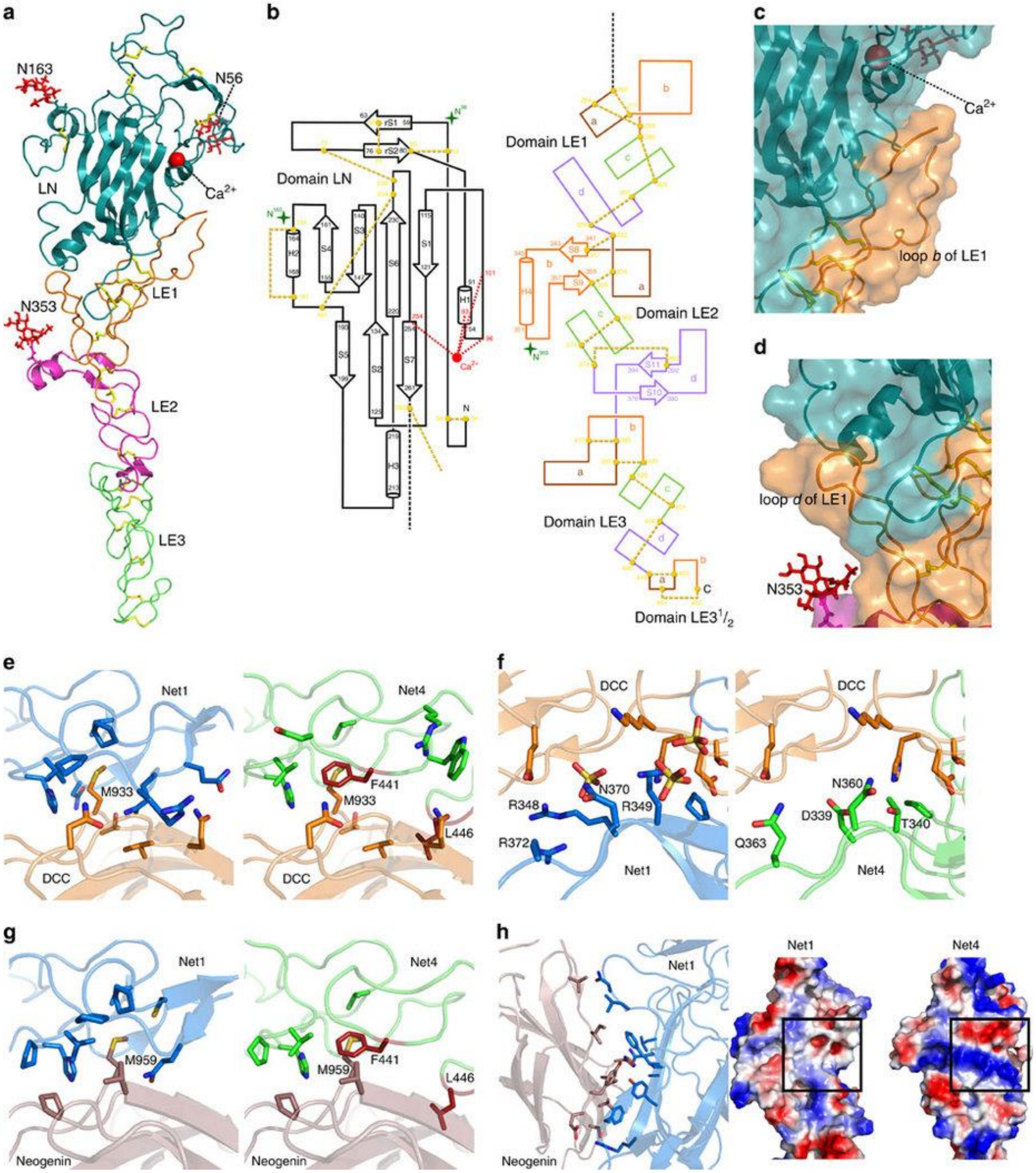


Figure 3-1. Overall Structure of Net4.

(**a**) Ribbon model of Net4-ΔC viewed in the “open face” orientation. The domains LN (teal), LE1 (orange) LE2 (magenta) and LE3-3$^{1}/_{2}$ (green) are colored accordingly. N-linked glycans are drawn as red sticks, disulfide bridges are drawn as yellow sticks, and the calcium ion as a red sphere. (**b**) Secondary structure diagram. Disulfide connectivity in domains LN and LE are indicated by yellow dotted lines, visible N-linked glycans by green stars with the Asn residues N^{56}, N^{163} and N^{353} highlighted. β-strands are depicted as arrows and the α-helices as cylinders, each labelled with the first and last residue. Domain LN contains 12 cysteines, which form a unique disulfide pattern in the order of 1-2, 3-7, 4-6, 5-12, 8-9 and 10-11. The cysteine residues 1-3, 2-4, 5-6, and 7-8 in subdomains LE1 to LE3 are bridged together forming the loop segments *a-d* accordingly. The individual loop segments are shown in different colour codes. (**c**) Interface between domains LE1 and LN. The unique loop *b* segment between Cys271 and Cys295 of subdomain LE1 is leaning towards the calcium-binding edge. The KAPG motif is free in solution and not visible in the structure. (**d**) Secondary interface between domains LE1 and LN. The loop *d* segment between Cys307 and Cys329 of subdomain LE1 is in contact with the N-terminus and the loop S5-H3 of domain LN. (**e**) Complex of Net1 (marine) (PDB 4OVE) and DCC (orange) (PDB 4URT) at site 1 with interacting residues shown at full opacity and the same complex with Net4 (green) replacing Net1. Though many interacting residues are conserved between the netrins, F441 of Net4 (red) occludes M933 from a hydrophobic binding pocket important for complex formation. Note also L456 (red) of Net4 clashes with DCC. (**f**) Complex of Net1 (marine) and DCC (orange) (PDB 4URT) at site 2 with interacting residues shown at full opacity and the same complex with Net4 (green) replacing Net1. Again, many interacting residues are conserved between the netrins, but a positive patch on Net1 composed of R348, R349, R351, and R372 that interacts with DCC both directly and indirectly through a sulfate ion is absent in Net4. (**g**) Complex of Net1 (marine) and neogenin (pink) (PDB 4PLN) at site 1 with interacting residues shown at full opacity and the same complex with Net4 (green) replacing Net1. Similar to site 1 of the DCC-Net1 complex, F441 (red) of Net4 occludes Met959 of neogenin from the hydrophobic pocket. L456 (red) of Net4 again clashes with the binding partner. (**h**) Net1 (marine) in complex with neogenin (pink) at site 2. The buried surface area of this interaction is ~680Å^{2}, and covers a large portion of the LN domain of Net1. None of the residues from Net1 that interact with neogenin are conserved in Net4. Note the difference in the surface charge at the neogenin binding site (indicated by a box) of Net1 (left) and the corresponding site on Net4 (right).

Net4 does not interact with Net1 receptors

As the DCC and UNC5B binding epitopes of Net1 are absent in the Net4 structure[22-24], we examined whether Net4 could interact with the ectodomains of DCC, UNC5B, A, C, D and more generally A2b, neogenin, as well as DSCAM that have been previously reported to interact with

Net1 using solid-phase binding assays and Octet technology. Structural analysis of Net4 and previously published Net1 complex structures reveals incompatibility between the binding residues of Net1 and equivalently positioned residues of Net4 (**Figure 3-1e-h**). For example, Phe441 of Net4 occludes the conserved Met from the hydrophobic pocket in both DCC and neogenin (**Figure 3-1e**, **g**), and Leu458 of loop *b* in the terminal LE domain of Net4 would preclude interaction with the netrin receptor. The absence of the positively charged residues of Net1 (Arg348, Arg349, Arg351, and Arg372) necessary for a sulfate-mediated interaction with DCC in Net4 would hinder interaction between DCC and Net4 (**Figure 3-1f**). The binding of neogenin to the LN domain involves 10 residues of Net1, which are completely non-conserved in Net4, as presented by the electrostatic surface potential of Figure 3-1h. Figure 3-2a reveals that while Net1 interacts in the nanomolar range with DCC and UNC5B, Net4 does not show any interaction with any of the proposed Net1 receptors including DCC/neogenin, UNC5 and integrins (**Figure 3-2a** and **Appendix 1, Figure I-3**). Sequence alignments of netrin family proteins further highlight that almost all residues involved in Net1 interaction with the canonical netrin receptors are not conserved in Net4 (**Appendix 1, Figure I-4**). Together these observations support the view that Net4 does not interact directly with the known Net1 receptors, though it is possible in some specific settings that a protein complex may include Net4 and Net1 receptors.

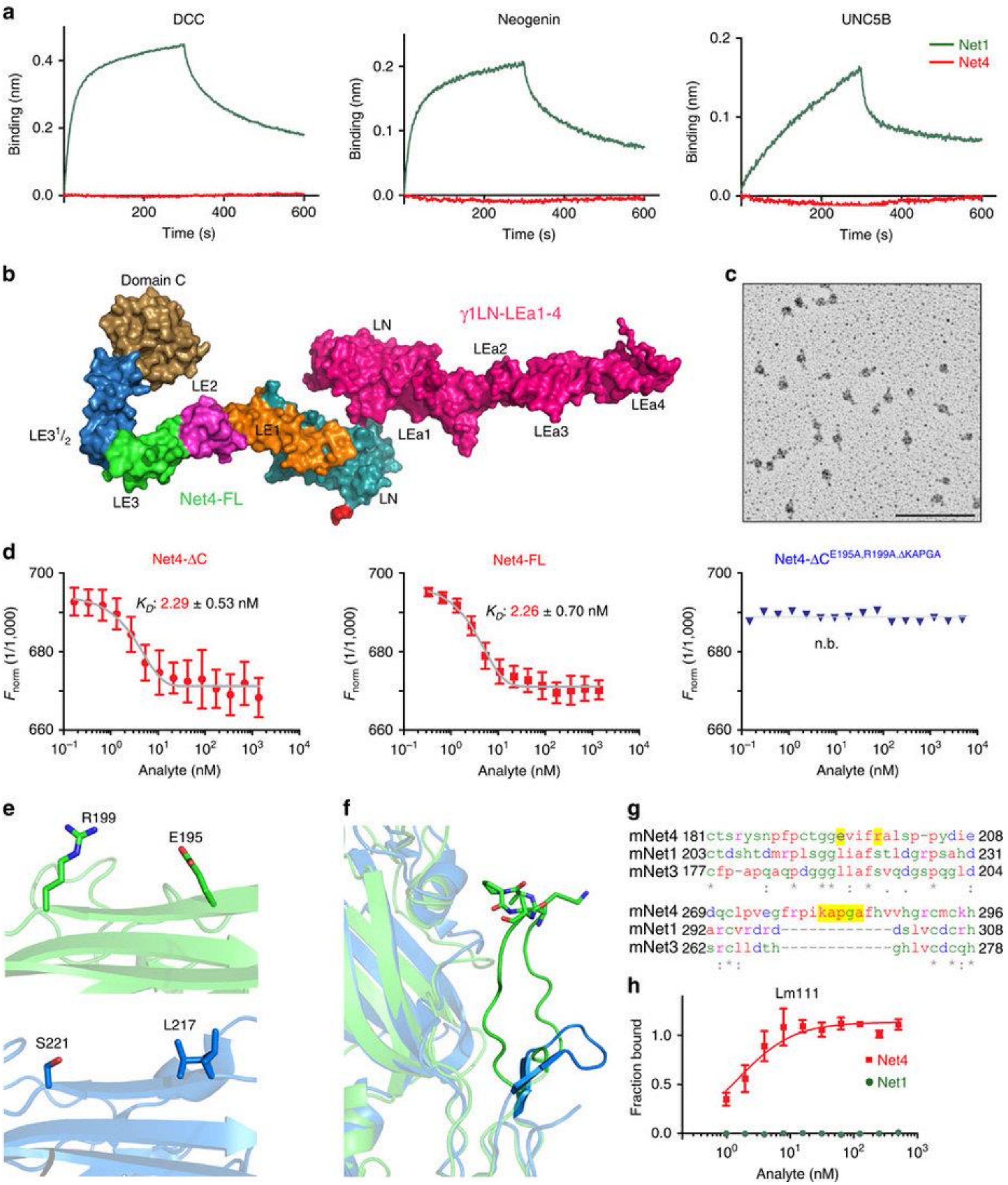


Figure 3-2. Binding comparison of Net4 and Net1 to netrin receptors and laminin.

(**a**) Biolayer interferometry binding kinetic analysis of DCC, neogenin, and UNC5B to Net4 (red) as well as Net1 (green). AHC sensors were coated with the Fc chimera proteins. Association (from 0 to 300 s) and dissociation (from 300 to 600 s) phases are shown for 50 nM of both netrins. (**b**) The SAXS model for the Net4-FL-γ1LN-LEa1-4 complex was generated by first calculating models for Net4-ΔC-γ1LN-LEa1-4 complex (**Appendix 1, Figure I-5d**) followed by the addition of the LE and globular domain C. This model highlights the role of N-terminal globular domains in the interaction. The agreement between experimentally collected SAXS data and model derived SAXS data is presented in Appendix 1, Table I-1. (**c**) Rotary-shadowing electron microscopy image of Net4-ΔC-γ1LN-LEa1-4 clearly showing a 1:1 complex mediated via the globular domains. (**d**) Microscale thermophoresis binding analysis of different Net4 mutants to labeled γ1LN-LEa1-4. Binding of Net4 lacking the netrin-specific C domain (Net4-ΔC). Binding of full-length Net4 (Net4-FL). Analysis of the binding of a Net4 mutant (Net4-$\Delta C^{\Delta KAPGA}$) in which the specific *loop b* (KAPGA) within LE1 was deleted. Binding of a combined Net4 mutant (Net4-$\Delta C^{E195A,R199A,\Delta KAPGA}$). Error bars, s.d. (*n* = 3 independent technical replicates). K_D values are shown in the graph (n.b., no binding). (**e** and **f**) Laminin-binding epitopes within Net4 (green) compared to the equivalent positions in Net1 (marine). Note loop *b* of LE1 of Net4 has an extensive contact area with the LN domain, and the KAPGA motif (shown in stick notation) is highly flexible. (**g**) Sequence alignment of Net4, Net1, and netrin-3 showing the laminin binding region. (**h**) Solid-phase binding study of Net4 and Net1 to immobilized laminin γ1 (γ1LN-LEa1-4). Error bars, s.d. (*n* = 3 independent technical replicates). Scale bar, (**c**) 100 nm.

Net4 forms a high affinity complex with laminin γ1

Although full-length Net4 is reported to exist as a dimer[3] possibly bridging two laminin heterotrimers, our crosslinking experiments as well as hydrodynamic studies reveal that full-length Net4 (Net4-FL) forms a high-affinity complex with laminin γ1 (Lmγ1) in a 1:1 molar ratio (**Appendix 1, Figure I-5a-e**). The hydrodynamic radius (r_H) from dynamic light scattering (DLS, 5.77 ± 0.02 nm) and the sedimentation coefficient from analytical ultracentrifugation (AUC, 5.15 ± 0.01 S) confirm unambiguously the existence of an equimolar heterodimeric complex with a molecular mass of 116 kDa. The rotary shadowing experiments reveal a central globular structure with two elongated rods oriented in opposite directions, consistent with a LN domain-mediated major binding interface with both elongated LE subdomains oriented in an antiparallel fashion (**Figure 3-2c**). Small Angle X-ray Scattering (SAXS) studies of γ1LN-LEa1-4 in complex with full-length Net4 (Net4-FL) yield an extended conformation with a maximum particle size (D_{max}) of 19 nm (**Table 1** and **Appendix 1, Figure I-5d**). Rigid body fitting allows the reconstruction of the assembly of full-length Net4 and laminin γ1 (Net4-FL-γ1LN-LEa1-4) in contact through the LN domains, with the C-terminal domain of Net4 (domain C) curling back over the LE domains. As presented in Figure 3-2b, both N-terminal β-sandwiches are oriented like the palms of two hands with the open face sheets S5-S2-S7 (Net4) and S6-S3-S8 (laminin γ1) in contact with each other (**Figure 3-2b**). The models were validated extensively by comparing experimentally determined hydrodynamic properties with model derived hydrodynamic properties as presented in Appendix 1, Table I-1. The comparison of Net4-ΔC and Net4-FL binding to laminin γ1 indicated equal affinities in the range of K_D = 2 nM, validating the exclusion of domain C of Net4 in complex formation (**Figure 3-2d**). To decipher individual amino acid residues responsible for Net4-laminin γ1 interactions, we designed mutants based on the three-dimensional structure of Net4 as well as the phylogenetic conservation (**Appendix 1, Figure I-1a, b**). A structure-guided deletion of the unique Net4 loop *b* (KAPGA) resulted in a 25-fold reduction of K_D from 2 nM (Net4-ΔC) to approximately 50 nM (Net4-ΔC$^{\Delta KAPGA}$) (**Appendix 1, Figure 3-6a**, **b**). Thus, Net4-ΔC$^{\Delta KAPGA}$ emphasizes the importance of the KAPGA loop in complex formation and gives a rationale for the lower affinity of laminin β1, which lacks the KAPGA motif, to laminin γ1 (K_D of 5 μM)[33,35]. Moreover, we identified two residues (E195 and R199) within the LN domain of Net4 involved in laminin γ1 binding. Mutation of these two residues results in a decreased K_D from 2 nM to 66 nM (**Appendix 1, Figure I-6c**). Additionally, mutation of E195 as well as R199 together with deletion of loop *b* (Net4-ΔC$^{E195A,R199A,\Delta KAPGA}$) completely abolishes binding in all performed binding assays (**Figure 3-2d** and **Appendix 1, Figure I-6a**, **b**). Cross-linking experiments confirm that full-length Net4 (Net4-FL) exists as a monomer and dimer, because increasing cross-linker concentration results in an increase in dimeric and multimeric Net4-FL (**Appendix 1, Figure I-5a**). The r_H of Net4-FL was determined to be 5.76 ± 0.02 nm which is very similar to that of Net4-FL-γ1LN-LEa1-4 (5.77 ± 0.02 nm) and Net4-ΔC-γ1LN-LEa1-4 (5.40 ± 0.10 nm) complexes, but significantly higher than Net4-ΔC (4.60 ± 0.20 nm), implying that full-length Net4 (Net4-FL) behaves as a mixture of monomer and dimer in solution (**Table 1** and **Appendix 1, Figure I-5e**). The maximum particle size (D_{max}) calculated from SAXS data for Net4-FL is significantly higher than the D_{max} of Net4-ΔC, but similar to both complexes. The AUC studies of Net4-FL provided two peaks corresponding to a monomer and to a dimer for Net4-FL, while Net4-ΔC[37] and complexes of Net4-FL and Net4-ΔC with γ1LN-LEa1-4 yield only a single peak (**Appendix 1, Figure I-5e**). These experiments provide evidence that full-length Net4 (Net4-FL) exists as a mixture of monomer and dimer in solution and that laminin γ1 is capable of destabilizing the Net4-FL dimer. Taken together, our results establish that full-length Net4 binds to laminin γ1 in a 1:1 stoichiometry. Additionally, biochemical studies identified different epitopes within Net4 responsible for laminin interaction. Our Net4 laminin-binding mutants can now be used to determine the specific role of the Net4-laminin complex in functional studies because these mutants do not affect cell adhesion properties of Net4 (**Appendix 1, Figure I-6d**) indicating their exclusive role in laminin binding. Moreover, comparing the structure of Net4 and Net1 reveals that the laminin-binding epitopes are not conserved within Net1 (**Figure 3-2e-g**) and indeed Net1 does not interact with laminin 111 (**Figure 3-2h**).

Net4 disrupts pre-existing laminin networks

Net4 has been proposed to block laminin polymerization[25]. However, due to its nanomolar binding to the laminin $\gamma1$ chain, we hypothesized that Net4 not only blocks laminin polymerization but also disrupts the pre-existing laminin network. First, we analyzed whether Net4 and the identified laminin-binding mutant (E195A, R199A), which still binds in a nanomolar range to the laminin $\gamma1$ chain, are able to block laminin polymerization (**Figure 3-3a** and **Appendix 1, Figure I-7a**). Interestingly, Net4-ΔC as well as the laminin-binding mutant Net4-$\Delta C^{E195A,R199A}$ blocked laminin polymerization to the same extent as the N-terminal fragments of laminin $\beta1$ and $\gamma1$ (**Figure 3-3b**). To address our hypothesis that Net4 is able to disrupt pre-existing laminin networks, we performed a comprehensive series of experiments, including (i) laminin network disruption (**Appendix 1, Figure I-7b**), (ii) size-exclusion chromatography, and (iii) atomic force microscopy of a polymerized Matrigel matrix. Net4 is able to disrupt pre-existing laminin 111 networks and the ternary node complex formed by the laminin $\alpha1$, $\beta1$, and $\gamma1$ chain (**Figure 3-3c-e**). Moreover, Net4 treatment dramatically weakens the stiffness of the Matrigel matrix (**Figure 3-3f**). Thus, all approaches clearly highlight that Net4 is able to disrupt pre-existing laminin networks. Though the laminin-binding mutant Net4-$\Delta C^{E195A,R199A}$ can block laminin polymerization, it is unable to disrupt pre-existing laminin networks (**Figure 3-3c**, **d**, **f**). Our results clearly demonstrate that N-terminal laminin fragments and Net4, as proposed previously[25,33] as well as the laminin-binding mutant, are able to block laminin polymerization due to their binding capacity to the N-terminal laminin (LN) domain (**Figure 3-3g**). These results establish that Net4 is able to disrupt pre-existing laminin networks in a non-enzymatic manner, whereas its respective laminin-binding mutant cannot (**Figure 3-3g**).

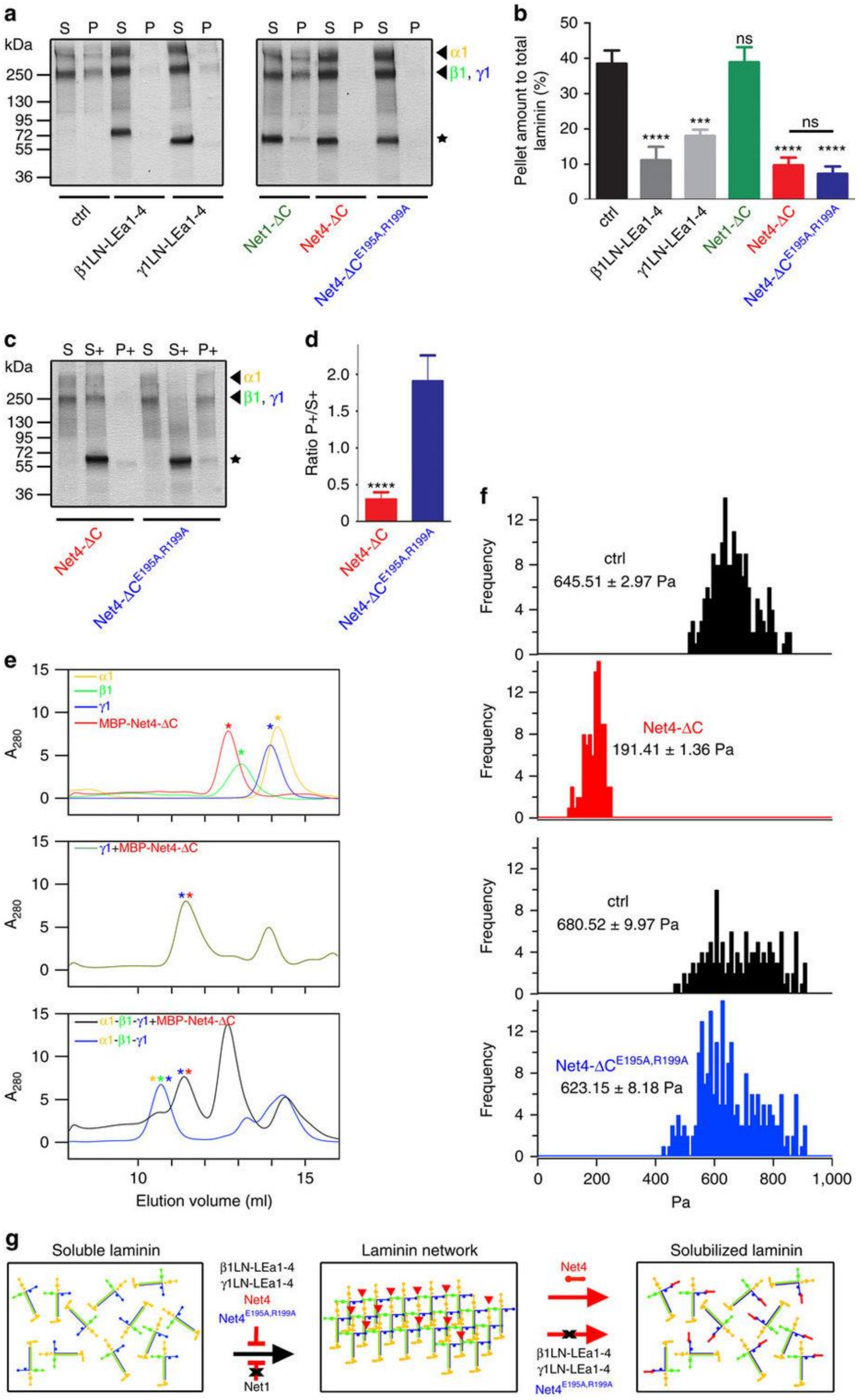

a
S P S P S P
S P S P S P
kDa
250
130
95
72
55
36
α1
β1, γ1
★
ctrl
β1LN-LEa1-4
γ1LN-LEa1-4
Net1-ΔC
Net4-ΔC
Net4-ΔC^{E195A,R199A}
b
Pellet amount to total laminin (%)
50
40
30
20
10
0
ns
****
***
ns
****
****
ctrl
β1LN-LEa1-4
γ1LN-LEa1-4
Net1-ΔC
Net4-ΔC
Net4-ΔC^{E195A,R199A}
c
S S+ P+ S S+ P+
kDa
250
130
95
72
55
36
α1
β1, γ1
★
Net4-ΔC
Net4-ΔC^{E195A,R199A}
d
Ratio P+/S+
2.0
1.5
1.0
0.5
0
****
Net4-ΔC
Net4-ΔC^{E195A,R199A}
e
A_{280}
15
10
5
0
α1
β1
γ1
MBP-Net4-ΔC
γ1+MBP-Net4-ΔC
α1-β1-γ1+MBP-Net4-ΔC
α1-β1-γ1
10
15
Elution volume (ml)
f
Frequency
12
8
4
0
ctrl
645.51 ± 2.97 Pa
Net4-ΔC
191.41 ± 1.36 Pa
ctrl
680.52 ± 9.97 Pa
Net4-ΔC^{E195A,R199A}
623.15 ± 8.18 Pa
0
200
400
600
800
1,000
Pa
g
Soluble laminin
β1LN-LEa1-4
γ1LN-LEa1-4
Net4
Net4^{E195A,R199A}
Net1
Laminin network
Net4
β1LN-LEa1-4
γ1LN-LEa1-4
Net4^{E195A,R199A}
Solubilized laminin

Figure 3-3. Net4 disrupts pre-existing laminin network

(**a**) Inhibition of laminin 111 polymerization (Lm111) by human laminin β1 (β1LN-LEa1-4), mouse laminin γ1 (γ1LN-LEa1-4), mouse Net1 (Net1-ΔC), mouse Net4-ΔC, and the laminin-binding mutant mouse Net4-ΔC^{E195A,R199A}. SDS-PAGE analysis of the pellet (P) and the supernatant fraction (S). The position of the laminin chains (arrow heads) and the recombinant proteins (stars) are indicated. (**b**) Densitometric analysis of three independent polymerization experiments and the percentage of the pellet fraction are displayed. (mean ± s.d.; *n* = 3; ****$P <$ 0.00006 (ctrl vs. β1LN-LEa1-4, Net4-ΔC, and Net4-ΔC^{E195A,R199A}) and ***$P = 0.0004$ (ctrl vs. γ1LN-LEa1-4); ns, not significant. *P*- values from two sided *t*-test). (**c**) The resulting fractions (supernatant of laminin polymerization (S), supernatant after addition of different proteins (S+), and polymer (P+)) were separated by SDS-PAGE, followed by Coomassie Brilliant-Blue staining. Different laminin chains are colour labelled (α1 (yellow), β1 (green), and γ1 (blue)) and added recombinant proteins are indicated with asterisks. (**d**) The ratio between polymer (P+) and supernatant (S+) was determined by densitometric analysis of at least three independent experiments (mean ± s.d.; *n* = 3; ****$P = 0.00005$). (**e**) N-terminal laminin fragments (LN-LEa1-4) of α1 (yellow), β1 (green) and γ1 (blue) as well as MBP-Net4-ΔC (red) single proteins were analyzed using analytical size-exclusion chromatography (top). SEC profile of the laminin fragment γ1LN-LEa1-4 (γ1) in complex with MBP-Net4-ΔC (green, middle). Complexes [α1-β1-γ1 (blue) and α1-β1-γ1 + MBP-Net4-ΔC (red)] were analyzed by SEC (bottom) revealing that the MBP-Net4-ΔC is able to disrupt the α1-β1-γ1 complex. Asterisks indicate the single proteins within the complex. (**f**) Atomic force microscopy anaylsis of a polymerized Matrigel matrix treated with Net4-ΔC or with the laminin-binding mutant Net4-ΔC^{E195A,R199A}. (**g**) Model depicting inhibition of laminin polymerization via laminin β1 (β1LN-LEa1-4), laminin γ1 (γ1LN-LEa1-4), mouse Net4-ΔC, and the laminin-binding mutant mouse Net4-ΔC^{E195A,R199A} but not via mouse Net1 (Net1-ΔC). The pre-existing laminin network can only be solubilized through Net4. Error bars, s.d. (*n* = 3 independent technical replicates).

Net4 interferes with BM assembly on the cell surface

As Net4 has a high-affinity interaction with laminin γ1, a key component of the BM, we investigated whether Net4 affects BM assembly. As an *in vitro* model we used Schwann cells, which are normally involved in the myelination of the peripheral nervous system. It has previously been shown that BM components assemble on Schwann cell surfaces only if laminin

111 heterotrimers interact via their LN domains[34]. Therefore, the stable heterodimeric complex of Net4 with laminin γ1 may prevent efficient BM assembly on Schwann cell surfaces similar to LN domain truncated laminin 111. Schwann cells were incubated with laminin 111, collagen IV, and nidogen-1, which normally results in a BM assembly on the cell surface[34]. Deposition of BM components was then monitored via immunofluorescence staining for collagen IV (Col-IV) and laminin (Lmγ1) (**Figure 3-4a**). First, laminin 111 binds to cell surface integrins via its C-terminal LG domains and assembles via the LN domains of the α1, β1, and γ1 chain. Both interactions are necessary for laminin deposition[34]. We observed strong laminin as well as collagen IV staining when treating cells with only laminin 111, collagen IV, and nidogen-1 (ctrl) (**Figure 3-4a**, **b**). After addition of wild-type Net4 (Net4-FL) together with BM components, we could neither detect laminin nor collagen IV deposition by immunofluorescence, whereas treatment with Net4-FL^{E195A,R199A} showed laminin as well as collagen IV stainings similar to the control situation (**Figure 3-4a**, **b**). Moreover, we analyzed the influence of different molar ratios of laminin 111 to Net4-FL as well as Net4-FL^{E195A,R199A} (1:0.4, 1:1.8, 1:7, and 1:28). Net4 significantly decreased laminin as well as collagen IV levels on the cell surface at ratios 1:7 and 1:28, whereas the laminin-binding mutant (Net4-FL^{E195A,R199A}) had no influence on laminin and collagen IV staining intensities (**Figure 3-4c**). In summary, these results indicate that Net4 can destabilize the polymerization of laminins containing the laminin γ1 chain followed by disassembly of the BM.

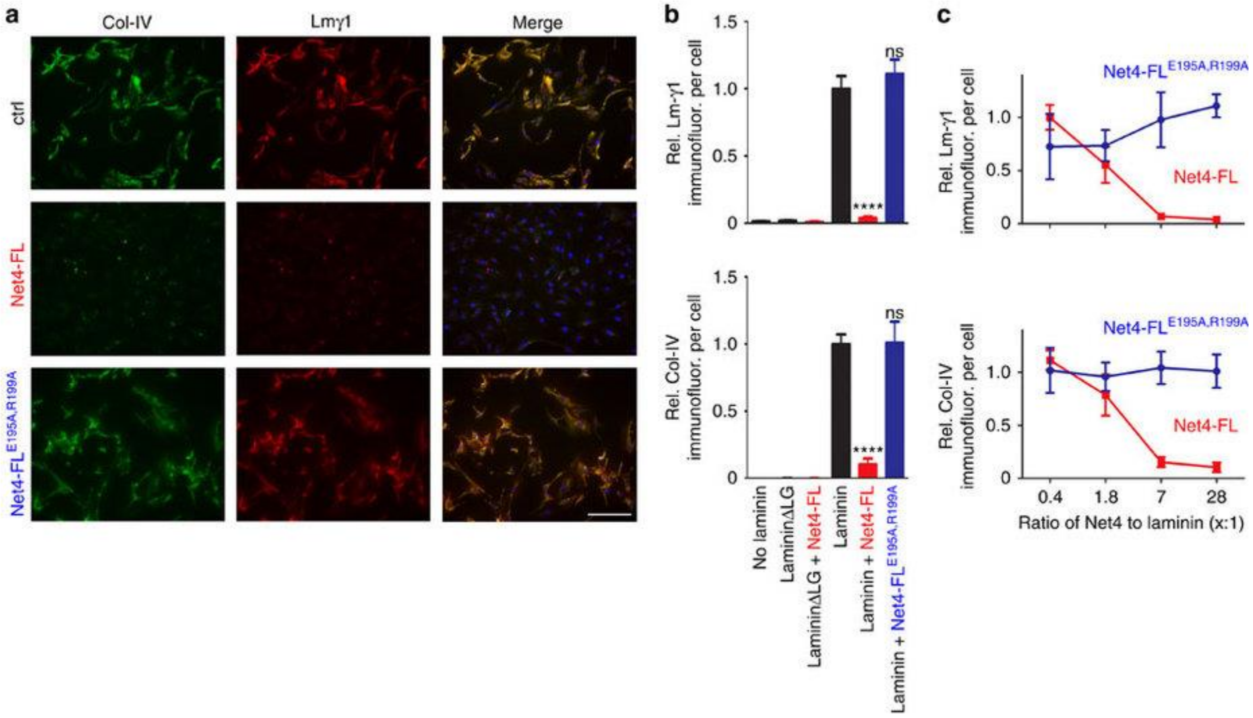


Figure 3-4. Net4 blocks laminin assembly on the cell surface.

(**a**) Within untreated (ctrl), Net4-FL-, and Net4-FL^{E195A,R199A}-treated (in 28-fold molar excess of laminin) Schwann cell extracellular matrix assembly assay BM components collagen IV (Col-IV, green) and laminin 111 (Lmγ1, red) were stained. Cells were counter-stained for DAPI (blue). (**b**) Staining intensities are displayed relative to the DAPI signal [relative Lmγ1 intensity/cell (left) and relative Col-IV intensity/cell, (right)] (mean ± s.d.; $n = 5$; ****$P < 0.00001$. *P*- values from two sided *t*-test). Error bars, s.d. ($n = 5$ independent cell cultures). *P* values were calculated by one-way Anova followed by a pairwise comparison using the Holm-Sidak method. (**c**) Concentration dependent effect of Net4-FL and Net4-FL^{E195A,R199A} on BM formation indicated through measurement of the relative Lmγ1 intensity/cell (left) and relative Col-IV intensity/cell (right). Net4 proteins were added together with the BM components laminin 111, collagen IV, and nidogen-1 with increasing molar ratios to laminin 111 (0.4, 1.8, 7, 28-fold excess). Scale bars, (**a**) 200 µm.

Net4 may affect axon outgrowth through matrix reorganization

Net4 has been postulated to play similar roles as Net1, mainly due to their conserved domain features[38-40]. Furthermore, it has been proposed that Net4 acts as an outgrowth promoting cue in *ex vivo* models such as the olfactory bulb neurite outgrowth[3]. We revisited this axon outgrowth function in the light of Net4's impact on the extracellular matrix. Olfactory bulb explants from E15 embryos were cultured in collagen I in the presence of the N-terminal laminin $\gamma 1$ fragment, Net4-ΔC, Net4-ΔC together with the N-terminal laminin $\gamma 1$ fragment, and the laminin-binding mutant Net4-$\Delta C^{E195A,R199A}$, which does not interact with laminin (**Figure 3-5a**). As shown in Figure 3-5b, while wild-type Net4 presence is associated with neurite outgrowth, the combination of Net4 with the N-terminal laminin $\gamma 1$ fragment as well as the laminin-binding mutant was unable to promote axon outgrowth (**Figure 3-5b**, **c**). These data support the view that, unlike Net1, the axon outgrowth activity of Net4 is not only due to guidance cue functionality *per se*. It may additionally be due to Net4 ability to disrupt the BM surrounding the olfactory bulb through interactions with laminin aiding axons in branching through this physical barrier, though we cannot exclude other interpretations. To further analyze this indirect function, we also explored the described role of Net4 on angiogenesis.

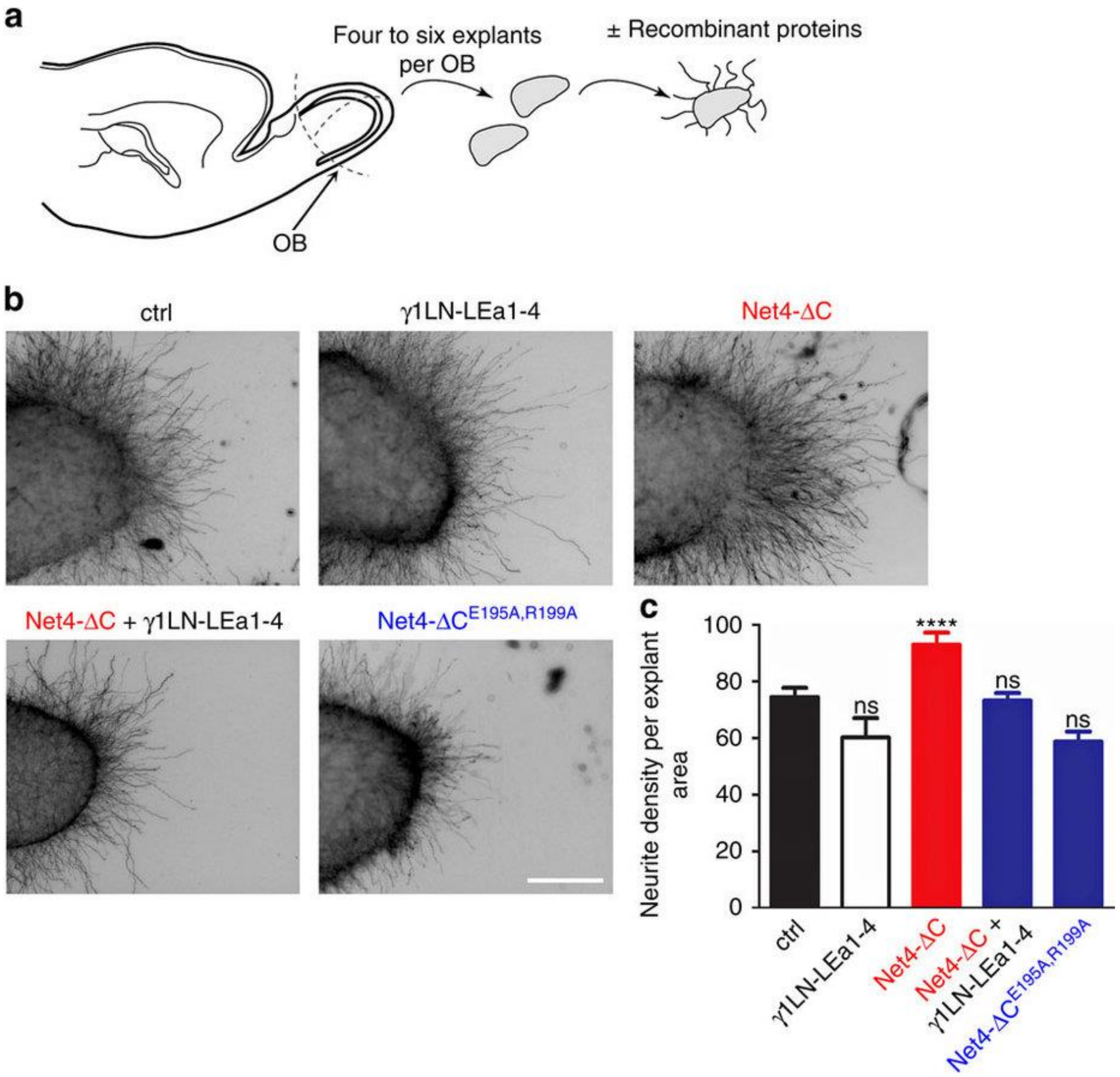


Figure 3-5. Net4 induces neurite outgrowth in a laminin dependent manner.

(**a**) Diagram depicting the olfactory bulb (OB) explants assay procedure. (**b**) Representative images of untreated (ctrl), laminin γ1 (γ1LN-LEa1-4)-, Net4 (Net4-ΔC)-, Net4 in combination with laminin γ1 (Net4-ΔC + γ1LN-LEa1-4)-, and Net4 laminin-binding mutant (Net4-$\Delta C^{E195A,R199A}$)-treated OB explants. (**c**) Quantification of neurite growth of OB explants (mean ±

s.e.m.; $n = 19$; $****P < 0.0001$, ns, not significant). Error bars, s.e.m. ($n = 19$ independent cell cultures). P values, two sided *t*-test. Scale bar (**b**) 200 µm.

Endothelial tubes require a functional basement membrane

Most studies investigating the role of Net4 in angiogenesis *in vitro* or *ex-vivo* are using artificial matrix for the culture *in vitro*[19,20,41-44]. The role of a functional laminin network within artificial matrix such as Matrigel for the formation of tube-like structures has so far not been addressed. A recent study revealed that Net4 inhibits the formation of endothelial tube-like structure at high concentrations (1000 nM); significantly higher than used in our neurite outgrowth assay[20]. This is difference in concentration dependence is likely due to the degree of basement membrane disruption required to observe the requisite effect. To confirm the results of the previous study, HDMECs were seeded on Matrigel and different concentrations (0, 200, and 1000 nM) of Net4 (Net4-ΔC) were added after cell attachment on Matrigel (**Appendix 1, Figure I-8a**). Untreated HDMECs formed cell clusters connected by honeycomb-like network structures after 14 h. Upon addition of 200 nM Net4 (Net4-ΔC), endothelial cells were able to form networks, whereas treatment with 1000 nM Net4 completely abolished tube-like structure formation (**Figure 3-6a** and **Appendix 1, Figure I-8a**, **b**). In addition, HDMEC tube-like structure formation was monitored for 14 h; upon Net4 treatment (1000 nM) cell morphology was dramatically changed and no network was established. Instead, cells moved together and formed pronounced aggregates within 5 h and remained clustered over the 14 h period (**Appendix 1, Figure I-8c**). On the basis of our hypothesis that Net4 disrupts laminin-laminin complexes, we further investigated whether Net4 alters tube-like structure formation in a cell-type independent manner.

Upon Net4 treatment, tube-like structure formation in microvascular (HDMEC), macrovascular (HUVEC), and lymphatic (HDLEC) endothelial cell lines was inhibited compared to control (**Appendix 1, Figure I-8d**). In addition, Net4 also destabilizes an already established vessel-like network on Matrigel 7 h post-treatment followed by increased cell aggregation after 14 h (**Appendix 1, Figure I-8e**). In a subsequent step, we further examined whether the exchange of the LN-LEa1 domains within the laminin γ1 chain fragment with the respective laminin-binding domains of Net4 LN-LE1 might lead to gain of function. Both chimeric fragments (chimera 2: Net4LN-LE1-2 - γ1LEa3-4 and chimera 3: Net4LN-LE1 - γ1LEa2-4) inhibited endothelial tube-like structure formation in the same way as Net4-ΔC, whereas the chimeric fragment (chimera 1: γ1LN-LEa1 - Net4LE2-3$^{1}/_{2}$) did not show any alteration of endothelial tube-like networks at the concentrations tested (**Appendix 1, Figure I-8f**).

To determine whether alterations in the laminin network are the trigger of Net4 mediated inhibition of tube-like structures, we performed the Matrigel assay with the structure-guided Net4 laminin-binding mutants (Net4-ΔC^{E195A}, Net4-ΔC^{R199A}, and Net4-ΔC$^{\Delta KAPGA}$, Net4-ΔC^{E195A,R199A}). HDMECs treated with these laminin γ1 binding mutants established cell clusters with filamentous interconnections in the same manner as untreated and laminin γ1LN-LEa1-4 fragment treated endothelial cells (**Figure 3-6a**, left; **Appendix 1, Figure I-9a**, **b**). Net4 significantly decreased cell cluster interconnections (tube number), whereas tube numbers in γ1LN-LEa1-4- and Net4 laminin-binding mutants-treated endothelial cells were equal to untreated endothelial cells (**Figure 3-6a**, right; **Appendix 1, Figure I-9a**, **b**). The inhibition of tube-like structure formation through Net4 was rescued via pre-incubation of Net4 together with the laminin-binding domain of the laminin γ1 chain (γ1LN-LEa1-4). Analysis of endothelial tube numbers showed no difference between untreated, treatment with γ1LN-LEa1-4, and combined treatment of γ1LN-LEa1-4 together with Net4 (**Figure 3-6b** and **Appendix 1, Figure I-9a**, **b**). This finding demonstrates that the laminin γ1 fragment competes with the interaction of Net4 to the laminin γ1 chain of the laminin 111 heterotrimer. Net4 can also block the formation of tube-like structures (**Appendix 1, Figure I-10a**, **b**) made by the mouse melanoma B16-F1 cell line[45], confirming the cell-independent activity of Net4 . To exclude that the observed activity of Net4 is only mediated when Net4 is deposited within the laminin network of the Matrigel matrix, we designed a Net4 protein fused to the N-terminal agrin domain (Agrin-Net4-ΔC). First, we tested the binding activity of Agrin-Net4-ΔC to the laminin γ1LN-LEa1-4 fragment and found it binds with the same affinity as wild-type Net4 (**Appendix 1, Figure I-6c**). Moreover, we tested the binding activity of the N-terminal agrin domain to laminin 111. Here, we took advantage of a previously generated laminin 111 construct lacking the N-terminal domains (LN-LEa1-4) of the laminin γ1 chain (Lm111Δγ1LN-LEa1-4)[34]. As expected, Net4 does not bind to Lm111Δγ1LN-LEa1-4 (**Appendix 1, Figure I-11a**), whereas Agrin-Net4-ΔC and the laminin-binding mutant fused to the N-terminal agrin domain (Agrin-Net4-ΔC^{E195A,R199A}) interact in the nanomolar range with Lm111Δγ1LN-LEa1-4 (**Appendix 1, Figure I-11b**). Treatment of HDMECs with Agrin-Net4-ΔC abolished tube-like structures to the same extend as wild-type Net4 (Net4-ΔC and Net4-FL, **Appendix 1, Figures I-9a** and **I-11c**). However, the laminin-binding mutant fused to the N-terminal agrin domain as well as the combination of Agrin-Net4-ΔC together with the laminin γ1LN-LEa1-4 fragment significantly blocked the activity of Net4 (**Appendix 1, Figure I-11c**).

Thus, our data support the view that a formed laminin network within the Matrigel matrix is essential to generate tube-like structures and that the reported inhibitory activity of Net4 is mediated through binding to the LN domain of the laminin γ1 chain and disruption of the BM.

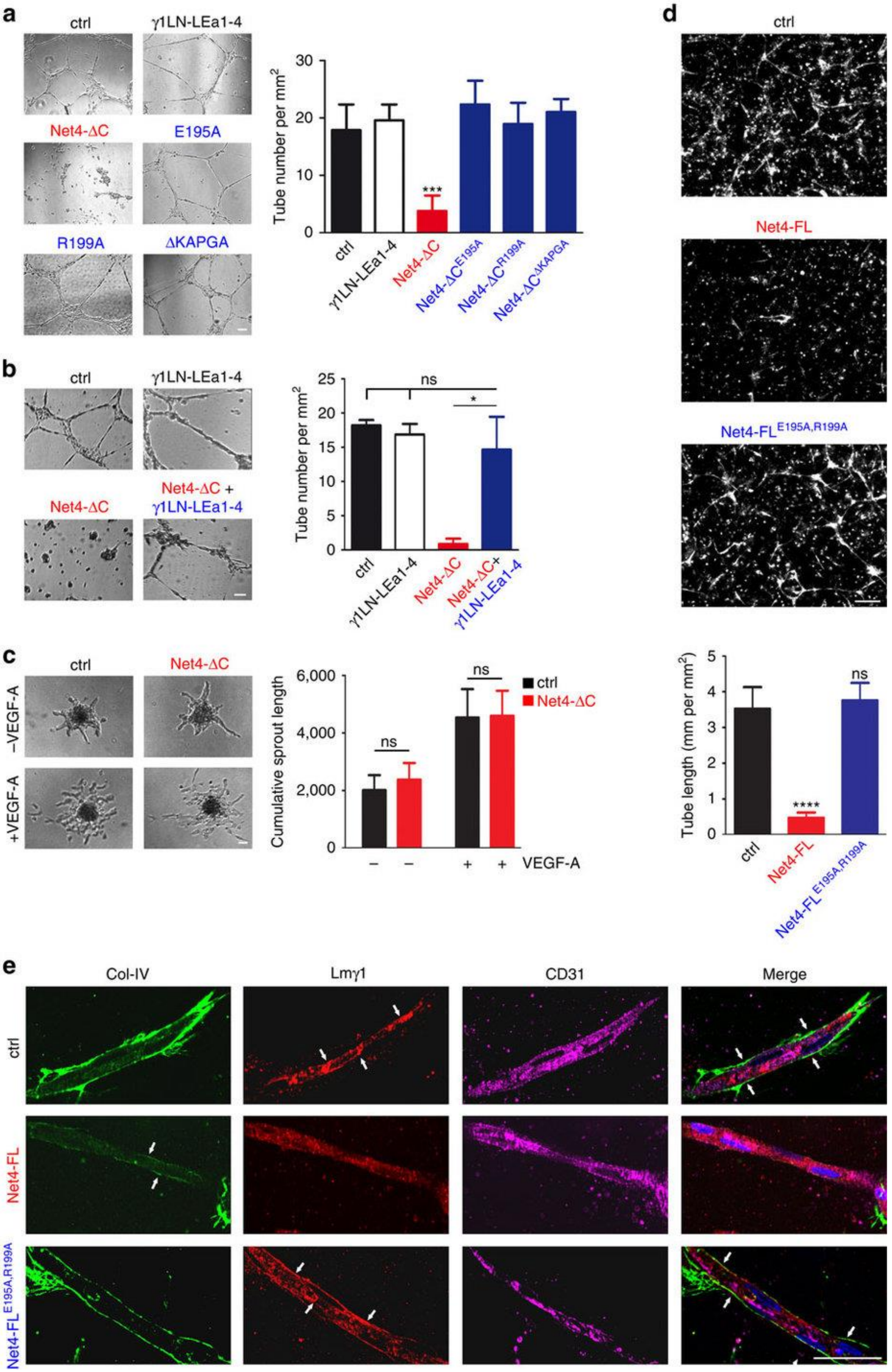

a
ctrl
γ1LN-LEa1-4
Net4-ΔC
E195A
R199A
ΔKAPGA
Tube number per mm²
30
20
10
0
***
ctrl
γ1LN-LEa1-4
Net4-ΔC
Net4-ΔC^E195A
Net4-ΔC^R199A
Net4-ΔC^ΔKAPGA
b
ctrl
γ1LN-LEa1-4
Net4-ΔC
Net4-ΔC +
γ1LN-LEa1-4
Tube number per mm²
25
20
15
10
5
0
ns
*
ctrl
γ1LN-LEa1-4
Net4-ΔC
Net4-ΔC+
γ1LN-LEa1-4
c
ctrl
Net4-ΔC
−VEGF-A
+VEGF-A
Cumulative sprout length
6,000
4,000
2,000
0
ns
ns
ctrl
Net4-ΔC
−
−
+
+
VEGF-A
d
ctrl
Net4-FL
Net4-FL^E195A,R199A
Tube length (mm per mm²)
5
4
3
2
1
0
ns
****
ctrl
Net4-FL
Net4-FL^E195A,R199A
e
Col-IV
Lmγ1
CD31
Merge
ctrl
Net4-FL
Net4-FL^E195A,R199A

Figure 3-6. Net4 activity on angiogenesis is dependent on Net4/laminin interaction.

(**a**) Analyses of tube-like structure inhibition by Net4-ΔC mutants (Net4-ΔC^{E195A}, Net4-ΔC^{R199A}, and Net4-ΔC$^{\Delta KAPGA}$) impaired in laminin γ1 binding (protein concentrations: 1 µM). Statistical analyses of tube formation via determining the number of cell cluster connections (tube number) (mean ± s.d.; *n* = 3; ****P* = 0.0007. *P*- values from two sided *t*-test) (right). Error bars, s.d. (*n* = 3 independent cell cultures). (**b**) Blocking of Net4-ΔC inhibited tube formation through equimolar addition of γ1LN-LEa1-4 (1 µM, molar ratio of Net4-ΔC:γ1LN-LEa1-4, 1:1). The laminin fragment γ1LN-LEa1-4 competes with Net4 binding to laminin (mean ± s.d.; *n* = 3; **P* = 0.035; ns, not significant. *P* values calculated by two sided *t*-test.). Error bars, s.d. (*n* = 3 independent cell cultures). (**c**) Treatment of VEGF-A induced spheroid sprouting of HDMECs embedded in collagen I by Net4-ΔC (1 µM). Analysis of sprouting events (ns, not significant). (**d**) Tube formation analyses of co-cultured endothelial and perivascular-like cells from untreated and Net4-FL- or Net4-FL^{E195AR199A}-treated cultures (30 nM). Representative images of CD31 staining are shown. The tube length [mm/mm^2] was quantified through the CD31 staining (mean ± s.d.; *n* = 5; *****P* < 0.00005). (**e**) Apotome images of Col-IV, Lmγ1, and CD31 stained tubes from control (ctrl), Net4-FL, and Net4-FL^{E195AR199A} treatment are displayed. Error bars, s.d. (*n* = 3, (**a**-**c**); *n* = 5, (**d**,**e**) independent cell cultures). Scale bars, (**a**-**d**) 100 µm; (**e**) 50 µm.

To unravel the physiological significance of our findings we performed a more physiologically relevant angiogenesis assay, where endothelial cells were co-cultured with pericytes to form tubes within a laminin-free collagen I matrix[28]. Here, endothelial cells and pericytes form a continuous basement membrane formed by laminin 511 and collagen IV. Treatment of the co-culture with Net4 completely abolished tube formation, whereas the laminin-binding mutant showed no effect (**Figure 3-6d**). Staining of endothelial tubes highlight that laminin as well as collagen IV is absent along the tubes in the Net4 treated set up (**Figure 3-6e**). These results not only suggest that Net4 disrupts the laminin network resulting in the decline of endothelial tubes but also emphasize that the laminin network is the key component of the vBM essential for vessel formation.

Net4 inhibits tumor growth through laminin binding

To determine how Net4 affects capillary networks an *ex ovo* chicken CAM angiogenesis assay[46] was performed. Established capillary networks were treated with full-length Net4 (Net4-FL) and the laminin-binding mutant (Net4-FL^{E195A,R199A}); blood and micro-capillaries were visualized using FITC-dextran injection 48 h post-treatment. In the untreated (ctrl) CAMs, a micro-vascular network appeared 48 h post-treatment. Upon addition of Net4, the capillary network dramatically changed as the observed vascularized area was significantly decreased (**Figure 3-7a**). Treating the CAMs with the laminin-binding mutant Net4-FL^{E195A,R199A} had no significant effect. Histological staining (hematoxylin, eosin; H&E) of CAMs to detect the capillary density of untreated and Net4-FL^{E195A,R199A}-treated CAMs displayed a dense capillary lining, whereas Net4-treated CAMs revealed fewer and more rounded capillaries (**Figure 3-7b**).

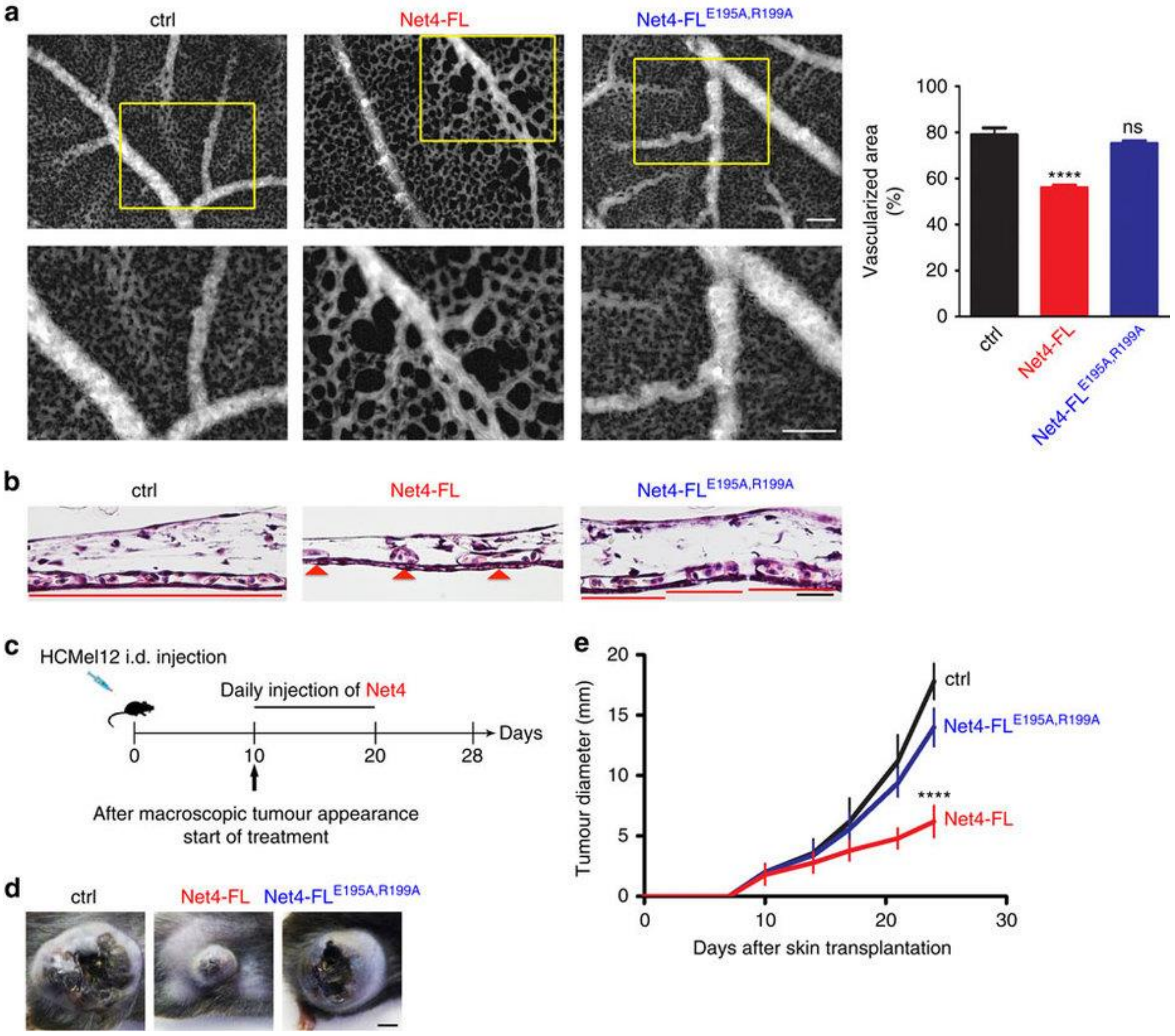
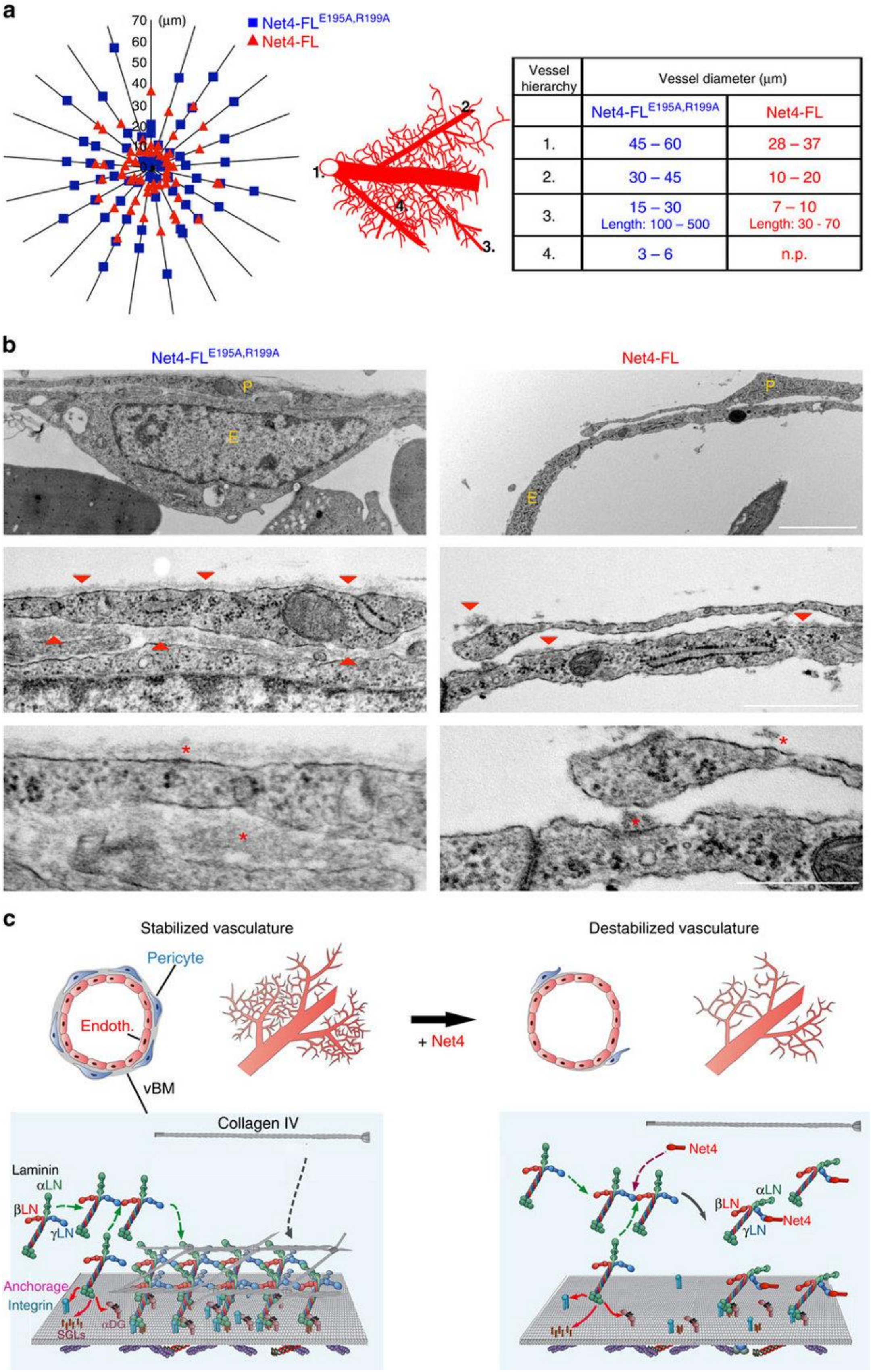


Figure 3-7. Laminin polymers are essential to maintain capillary networks and tumor progression.

(**a**) Visualization of capillaries via FITC-dextran injection of untreated (ctrl), Net4-FL-, and Net4-FL^{E195A,R199A}-treated CAMs after 48 h. The vascularized area from different protein treatments (left) was determined (see Methods) at 48 h (mean ± s.d.; $n = 5$; ****$P < 0.0001$. P-values from two sided *t*-test). Error bars, s.d. ($n = 5$ independent cell cultures). (**b**) H&E staining of untreated (ctrl) and Net4-FL-, and Net4-FL^{E195A,R199A}-treated CAMs after 48 h. Red lines and red arrow heads indicate the capillary density (capillaries are labelled with C). (**c**) Diagram showing the tumor treatment regimen. (**d**) Macroscopic images of melanoma treated daily with 1 µM of control protein (mouse serum albumin, left), Net4-FL (middle), Net4-FL^{E195A,R199A} (right). (**e**) Progression of melanoma treated with Net4-FL and the laminin γ1 binding mutant Net4-FL^{E195A,R199A} (mean ± s.d.; $n = 5$; $^{****}P < 0.0001$). Error bars, s.d. ($n = 5$ animals, female C57BL/6). *P* values, two sided *t*-test. Scale bars, (**a**) 100 µm; (**b**) 20 µm; (**d**) 5 mm.

One of the key pathophysiological roles of Net4 associated with its angiogenic activity is its inhibitory effect on tumor progression. Subcutaneously injected cancer cells overexpressing Net4 in mice establish tumors with fewer vessels than tumors established by untransfected cells[18,19]. Therefore, we investigated whether the tumor-suppressive activity of Net4 is mediated through binding to laminin. HCMel12 mouse melanoma cells were injected in mice and, upon macroscopic detection, tumor growth was monitored and daily injection of either a control protein (mouse serum albumin), Net4 or its laminin-binding mutant was administered (**Figure 3-7c**). The size of tumors treated with Net4 was reduced by 50% of the size in comparison to the control-treated animals. Moreover, our data establishes that the interaction of Net4 with laminin is essential for its tumor-suppressive activity as the laminin binding mutant (Net4-FL^{E195A,R199A}) had no impact on tumor growth (**Figure 3-7d**, **e**).

We then reexamined the *ex ovo* CAM model to further dissect the impact of Net4 on capillaries. In CAMs treated with Net4 or its laminin-binding mutant, the branching pattern is highly altered only in the Net4 treated CAM. Radial graph analysis of the capillary branching hierarchy clearly indicates that diameters of vessels within Net4-FL^{E195A,R199A}-treated CAMs span the full range of diameters found in an established vessel hierarchy. In contrast, vessels in Net4-FL-treated CAMs showed an altered distribution of diameters as large vessels exhibited smaller diameters and small capillaries were not present (**Figure 3-8a**, left). Net4-FL^{E195A,R199A}-treated CAMs display a distinct established hierarchy pattern, starting with large vessels (1., 45-60 µm diameter) from which smaller capillaries branch with decreasing diameter (2., 30-45 µm; 3., 15-30 µm; and 4., 3-6 µm). The altered vessel hierarchy in Net4-treated CAMs is clearly demonstrated by decreased diameters as large vessels (1., 28-37 µm) as well as small capillaries (2., 10-20 µm and 3., 7-10 µm) exhibit reduced calibers (**Figure 3-8a**, right). Thus, the laminin-laminin interaction is required to maintain proper capillary networks *in vivo*.



Vessel hierarchy	Vessel diameter (μm)	
	Net4-FL^{E195A,R199A}	Net4-FL
1.	45 – 60	28 – 37
2.	30 – 45	10 – 20
3.	15 – 30 Length: 100 – 500	7 – 10 Length: 30 - 70
4.	3 – 6	n.p.

Figure 3-8. Impact of Net4 on disruption of the laminin network and its effect on vascularization.

(**a**) Diagram indicates vessel diameter distribution (left) within Net4-FL^{E195A,R199A}-treated CAMs (blue squares), and the Net4-FL group (red triangle). Model showing the common established capillary hierarchy formed within the CAM. Numbers indicate sizes of different vessels in the common established hierarchy (middle). The table displays the diameter of different hierarchical vessels from Net4-FL^{E195A,R199A}- (blue) and Net4-FL-treated (red) experiments (right) (n.p., not present). (**b**) Ultrastructural analyses of Net4-FL^{E195A,R199A}- (blue) and Net4-FL-treated (red) CAMs using transmission electron microscopy. Images show the capillary endothelium surrounded by perivascular cells. The endothelium E is shown and perivascular cells (PVCs) are indicated as P. Electron microscopy images highlight the vascular basement membrane (vBM, red) arrow heads (middle) and red asterisks (bottom) between the endothelium and PVCs, which are surrounded by a vBM layer. (**c**) Model of Net4 effect on capillary cell types (endothelium (Endoth.) and pericyte (Pericyte)) and the vBM. Scale bars, (**b**) top: 2 µm; middle: 1 µm; bottom: 0.5 µm.

As Net4 appears to disconnect the laminin network *ex ovo*, we hypothesized that the disruption of the laminin network may affect the whole vBM thereby influencing the interaction of pericytes with the endothelium. Therefore, endothelial and perivascular cell distribution was studied in CAMs treated with the laminin-binding mutant (Net4-FL^{E195A,R199A}) or Net4. Ultrastructural analyses of Net4-FL^{E195A,R199A}-treated CAMs suggested that endothelial cells are closely surrounded by pericytes, whereas Net4 treatment leads to detachment of pericytes from the endothelium with altered pericyte morphology (**Figure 3-8b**, top). Embedding of pericytes into the vBM stabilizes newly-formed vessels. Net4-FL^{E195A,R199A}-treated CAMs exhibited a continuous vBM along the endothelium and surrounding pericytes. Upon Net4 treatment, the vBM almost disappeared on endothelial cell surfaces as well as along pericytes (**Figure 3-8b**, middle and bottom). These results indicate that Net4 dissociates laminin heterotrimers from the vBM and disturbs the endothelial/perivascular cell interaction *in vivo*.

Discussion

Far from its expected role as a guidance cue, a combination of structural, *in vitro* and *in vivo* experiments demonstrate that Net4 behaves in a completely different manner than its well-studied family member Net1. Our data demonstrate that Net4 not only strongly binds to laminin γ1 and blocks laminin polymerization, but also actively disrupts pre-existing laminin networks in a non-enzymatic manner. Net4 is clearly a unique member among the netrin family, showing non-enzymatic disruptive forces in matured basement membranes.

It has been reported that Net4 demonstrates axonal growth activity[3] and anti-angiogenic activity *in vitro* by inhibiting endothelial tube-like structure formation[18-20]. Furthermore, Net4 overexpression in different cancer cells reveals a decreased vascularization *in vivo*[18,19]. Though we cannot exclude an alternative function via unknown receptor, we propose that the basis of the effects reported on axonal growth, angiogenesis, and tumor progression by Net4 is due to the disruption of pre-existing laminin networks resulting in the disruption of the entire BM. Endothelial assemblies need to form a continuous vBM together with pericytes[28]. Our treatment studies with Net4 confirm this, as the disruption of the vBM results in the decline of endothelial tubes. In terms of the effect of Net4 on neurite outgrowth in the olfactory bulb assay the role of the BM is different. Here, axons need to branch through the BM as a physical barrier and the disruption of this barrier with Net4 supports axonal outgrowth. Interestingly, most of the cellular effects reported on Net4 have been performed using an artificial matrix such as Matrigel that contains mainly laminin 111[19,20,47,48]. As the LN domain of the γ1 chain is the major binding partner of Net4, we hypothesized that it affects the laminin network within the Matrigel matrix instead of directly affecting cell-signaling pathways. Similarly, inhibition of blood vessel formation by Net4 is believed to be triggered through binding to the Net1 receptors DCC (DCC, neogenin) and UNC5 (UNC5A-D)[18,19]. However, we could not detect direct interactions of Net4 with any of these receptors. Comparison of the high-resolution structures of Net1[22-24] and Net4 explains in detail this absence of interaction (**Figure 3-1e-h** and **Appendix 1, Figure I-1a**).

Structure based mutations of Net4 in combination with *in vivo* and *ex vivo* assays allowed assignment of Net4 functionality to its laminin-binding capability rather than Net4's interaction with Net1 receptors. Thus, our hypothesis is that Net4 is not a classic ligand activating a membrane receptor with a downstream signaling but rather a regulator of basal membrane maintenance, which indirectly impacts key functions such as axonal growth and angiogenesis. Our study supports the view that Net4 is a unique protein able to disrupt laminin γ1 containing BMs, and this could act as a useful tool in investigating the pivotal role of laminin-laminin interactions in biological processes, e.g., muscle formation, somitogenesis, and neural tube formation.

In the vasculature, laminin disassembly destabilizes three-dimensional endothelial tubes in an endothelial/pericyte-like cell co-culture via interference with the formation of a continuous BM along endothelial cell surfaces. Vessel maturation is affected by establishing a continuous BM along endothelial tube-linings. Cooperation between endothelial (ECs) and perivascular cells (PVCs) induces the production of pivotal matrix components such as nidogen-1, linking the collagen IV structure to the laminin network, and the α5 chain shaping laminin 511, which is the predominant laminin heterotrimer in mature capillaries[28]. The finding that treatment of an EC/PVC co-culture and chicken CAMs with Net4 results in decreased vascularized areas further highlights the essential role of vascular laminin networks in angiogenic processes. Furthermore, the Net4 laminin-binding mutant showed no effect on angiogenesis. Remarkably, Net4 disrupts the entire vBM lying between the endothelium and pericytes as well as the BM surrounding pericytes inhibiting stabilization of the capillary network and formation of the common vessel hierarchy *ex ovo*. Mural cells embedded in the vBM are defined as pericytes[29], and upon vBM loss, as demonstrated by ultrastructure analyses of Net4-treated CAMs, the phenotype of pericytes is altered to a more mural cell character. These findings are of potential importance in modulation of cancer progression where blood vessels are not always well organized, but appear destabilized and are often leaky, which hampers drug delivery and increases tumor cell spread[49-51]. The complex between laminin γ1 and Net4 clearly hinders the formation of, or even disrupts, the laminin network. Interestingly, disruption of the vBM leads to detachment of pericytes from the endothelium resulting in the collapse of the entire capillary network (Model pictured in **Figure 3-8c**), which may also help to reduce tumor progression. A recent study showed increased permeability in lymphatic vessels upon Net4 overexpression[21], which fits with the proposed model. Of interest, this model supports the view that the impact of Net4 depends highly on its ability to “titrate” laminin and thus an important aspect that remains to be investigated further in animal models is how physiological levels of Net4 actually modulates basal membranes and how in pathological conditions, up- or down-regulation of Net4 directly affects the BM and thus the cell behavior in this Net4 high/low environment. Several studies have shown that Net4 expression is modulated in cancer and as an example is inversely associated with cancer progression[52] suggesting that Net4 could represent an important modulator of tumoral matrix and thus a putative therapeutic target. Finally, our study not only highlight the key role of laminin networks within the vBM but also shows that Net4 disrupts pre-existing laminin networks and BMs in a non-enzymatic manner.

Tables

Table 3-1. Data collection and refinement statistics

	Net4-ΔC
Data collection	
Space group	C121
Cell dimensions	
a, *b*, *c* (Å)	107.755, 74.049, 74.672
α, β, γ (°)	90, 96.072, 90
Resolution (Å)	46.00-3.07 (3.12-3.07)†
No. reflections	39327 (1952)
R_{merge}	.194 (.776)
$I / \sigma I$	6.10 (2.0)
Completeness (%)	99.9 (99.8)
Redundancy	3.5 (3.5)
Refinement	
$R_{\mathrm{work}} / R_{\mathrm{free}}$	.2198/ .2626
No. atoms	3382
Protein	3259
Ligand/ion	96
Water	27
B-factors	45.90
Protein	44.10
Ligand/ion	110.30
Water	32.60
R.m.s. deviations	
Bond lengths (Å)	.003
Bond angles (°)	.77
Ramachandran Statistics	
Favoured (%)	92.2
Disallowed (%)	0.0

†Values in parentheses are for highest-resolution shell.

Methods

Recombinant protein expression and purification

Net4-FL, γ1LN-LEa1-4, and serum albumin from *Mus musculus,* (Net4: NP_067295, aa: 20-628, γ1LN-LEa1-4: NP_034813, aa: 34-492, serum albumin: NP_033784, aa: 1-608) and α1LN-LEa1-4, β1LN-LEa1-4 from *Homo sapiens* (α1LN-LEa1-4: NP_005550, aa: 18-453, β1LN-LEa1-4: NP_002282, aa: 31-509) were cloned into a modified pCEP vector with an N-terminal double *Strep* II-tag. Net4-ΔC from *Mus musculus* (NP_067295, aa: 20-462) was cloned into a modified pCEP vector with an N-terminal octa-histidine (His_8) tag or with an N-terminal double *Strep* II-tag. HEK293 cells were stably transfected followed by screening for subclones with a high level of protein expression. Net4-ΔC, Net4-FL, γ1LN-LEa1-4, α1LN-LEa1-4, and β1LN-LEa1-4 were purified by streptactin-sepharose (IBA) and Net4-ΔC by metal-affinity chromatography using nickel TALON beads (Clontech) followed by the removal of the double *Strep* II-tag and His_8-tag by thrombin digestion. The purified proteins were then dialyzed against 1 x PBS. Affinity purified proteins were further processed by size-exclusion chromatography (SEC) in TBS buffer (50 mM Tris-HCl, pH 7.5; 200 mM NaCl) on a Superdex 200 column using the ÄKTA FPLC system (GE Healthcare Life Sciences USA) equipped with UV absorbance detector and UNICORN v5.11 software as previously described. The complex of Net4-FL-γ1LN-LEa1-4 and Net4-ΔC-γ1LN-LEa1-4 was purified using SEC by mixing the peak fractions for individual species. Protein concentrations were measured using extinction coefficient and molecular weight values derived from the ProtParam utility[53] available on the ExPASy server[54].

Design of chimeric constructs and site-directed mutagenesis

To design Net4 containing LNγ1 binding domains we generated chimeric molecules by overlap PCR using the Q5 (New England Biolabs) polymerase. The domains from Net4 (Net4: NP_067295) were replaced by the corresponding laminin γ1 sequences (laminin γ1: NP_034813). The following constructs were amplified, cloned (with an N-terminal double *Strep* 2-tag), verified by sequencing, expressed, and purified: Net4LN-LE1-2 - γ1LEa3-4 (Net4 aa: 20-394 fused to laminin γ1 aa: 396-492), Net4LN-LE1 - γ1LEa2-4 (Net4 aa: 20-331 fused to laminin γ1 aa: 340-492), γ1LN-LEa1 - Net4LE2-3$^1/_2$ (Laminin γ1 aa: 44-339 fused to Net4 aa: 332-462), and Net4-ΔC - N-term agrin (NP_067295, aa: 20-462, E195A, R199A, fused to NP_067617, aa: 33 – 164). Furthermore, the following mutated Net4 molecules were generated: Net4-FL: (NP_067295, aa: 20-628, E195AR199A), Net4-ΔC$^{E195A,R199A,\Delta KAPGA}$: (NP_067295, aa: 20-462, E195A, R199A, and Δ281KAPGA285), Net4-ΔC^{E195A,R199A}: (NP_067295, aa: 20-462, E195A, R199A, and Δ281KAPGA285), and Net4-ΔC^{E195A,R199A} - N-term agrin (NP_067295, aa: 20-462, E195A, R199A, fused to NP_067617, aa: 33 - 164). All produced proteins are shown in **Appendix 1, Figure I-11**.

Structure determination of Net4

Net4-ΔC was applied to a Superdex 200 SEC column (GE Healthcare) equilibrated with 50 mMTris/Tris-HCl, pH 7.5, 0.2 M NaCl at room temperature and the collected peak fractions were immediately concentrated to 1.2 mg ml^{-1} and used for crystallization. Net4-ΔC crystals were grown by hanging-drop vapour diffusion at 293 K by mixing 2 µL of protein solution, 1.6 µl of reservoir solution (20% PEG3350, 0.2 M calcium acetate, and 0.1 M sodium cacodylate pH 6.5), and 0.4 µl of 1 M NDSB-256 from the Hampton HR2-428 additive screen (Hampton Research). Crystals appeared after 1 week and were soaked in mother liquor containing 10% ethylene glycol for 5-10 min before flash freezing in liquid nitrogen. Diffraction data was collected on a Rigaku MM-007HF (λ=1.54178Å) at 100 K. The dataset was indexed, integrated and scaled with the HKL2000 suite. The phases were determined by molecular replacement using a polyserine laminin β homology model (PDB 4AQS[55]) from Swissmodel in Phaser[56]. The model was built and refined using Coot[57], the Phenix Refine and Autobuild programs[58], and Refmac[59,60].

Octet binding studies

Binding of Net4 or Net1 to DCC, neogenin, and UNC5H2 was analyzed by biolayer interferometry using the Octet RED96 system (Pall ForteBio). Recombinant human DCC/Fc, mouse neogenin/Fc, and rat UNC5H2/Fc chimera were obtained from Bio-Techne. Proteins were stored frozen and diluted into binding buffer (PBS, 0.1% BSA, 0.02% Tween 20). Binding assays were performed in 96-well microtiter plates at 30°C with orbital sensor agitation at 1000 rpm. Fc chimera proteins (2 µg ml^{-1}) were loaded onto anti-human Fc (AHC) biosensors for 10 min. Biosensors were washed with binding buffer for 60 s and placed in wells containing the various netrins (50 nM in binding buffer). Association was observed for 10 min, then biosensors are incubated in binding buffer for additional 10 min to observe dissociation of the complex. Uncoated biosensors were run in parallel for non-specific binding. Binding curves were analyzed by Octet User software (ForteBio). Three independent technical replicates were performed.

Microscale thermophoresis binding assay

Binding of the labeled Laminin γ1 short arm fragment (γ1LN-LEa1-4) to Net4 variants and its respective mutants were measured using microscale thermophoresis. A range of concentrations of the required ligands (from 0.08 to 1000 nM for Net4 and from 0.45 to 4000 nM for Net4 mutants) were incubated with 2 nM of γ1LN-LEa1-4 in the assay buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$ supplemented with 3% BSA) for 10 min. Samples were loaded into NanoTemper glass capillaries and microscale thermophoresis was carried out using 10% LED power and 80% MST power on a NanoTemper monolith NT.115[61]. K_D values were calculated from independent technical triplicate experiments using the mass action equation via the NanoTemper software.

Laminin network disruption assay

The laminin polymerization blocking assay was performed as described in the Appendix 1, Figure I-9b and the laminin polymer disruption assay set up is described in the Appendix 1, Figure I-9c. Mouse laminin 111 (Sigma-Aldrich, Germany) was diluted to 0.35 µM in 50 mM Tis-HCl, pH 7.4; 150 mM NaCl; 0.1% Triton-X100; 1mM $CaCl_2$ and incubated at 37°C for 3 h to polymerize. Afterwards, polymerized laminin was centrifuged (10,000 g; 15 min) to separate formed laminin polymer from non-polymeric laminin. Different proteins (Net4-ΔC, Net4-ΔC^{E195A,R199A}) were added at a concentration of 2.1 µM respectively to the laminin polymer at 37°C for an additional 3 h. Samples were centrifuged and all three fractions (supernatant (S), supernatant+ (S+), pellet+ (P+)) were analyzed by SDS-PAGE followed by fast Coomassie Blue staining (Thermo Scientific, Germany). Band intensities were determined using the ImageJ software and ratios between the polymer fraction (P+) and supernatant (S+) are displayed.

Laminin assembly on Schwann cell surface

Rat Schwann cells were cultured in Dulbecco's Eagle medium, 10% fetal calf serum (Atlanta Biological), neuregulin (0.5 µg ml^{-1}, Sigma), forskalin (0.2 µg ml^{-1}, Sigma) and penicillin-streptomycin. Cells at passage 22-29 were seeded onto 24-well dishes (Sigma) and incubated with collagen IV, nidogen-1, and different laminin constructs. Additionally, cells were treated with 28-fold excess compared to laminin of Net4-FL, Net4-FL^{E195A,R199A} at 37^{o}C for 1 h. Schwann cells were rinsed with phosphate-buffered saline (PBS) and fixed in 3%

paraformaldehyde for 20 minutes followed by blocking with 5% goat serum in PBS overnight at 4ºC. Cells were incubated with anti-laminin γ1 (1:200, Millipore, MAB1920) and anti-collagen IV (10 µg/mL, Millipore, AB756P) at room temperature for 1 h. Alexa Fluor 647 goat anti-mouse and Alexa Fluor 488 goat anti-rabbit secondary antibodies (Molecular Probes) were used at 1:100 and counterstained with 4',6-diamindino-2-phenylindol. Digital images of protein immunofluorescene levels were recorded with IPLAB 3.5 software (Scanalytics) and quantified with ImageJ. Validation information for commercial antibodies is available on the manufacturers' websites. All experiments were performed in groups of five independent cultures and repeated at least twice.

Neurite outgrowth assay

Olfactory bulb (OB) explants were isolated from rat embryos E15 (*n* = 19 Embryos at E15 were collected from a pregnant Sprague Dawley female rat). OB explants were embedded in Collagen I isolated from rat tail and cultured with 28 nM of Net4-FL or Net4-FL^{E195A,R199A} for 24 h. Explants were then fixed in 4% PFA at room temperature for 1h and permeabilized with PBST (PBS, 0.1% Triton-X100) for 10 min. Next, explants were blocked with PBST, 1% normal donkey serum (Jackson immunoresearch) at room temperature for 2 h and incubated with mouse polyclonal Tuj1 antibody, 1/500 (babco, MMS435P) diluted in PBST at 4°C overnight. After washing 3 times with PBST, explants were incubated with Alexa488 donkey anti mouse (1/500; Life Technologies, A21202) at room temperature for 1 h, followed by 3 washes in PBST and mount with mowiol. Explants were imaged with an Axiovert200M light microscope (Zeiss).

Neurite and explant area were analysed with FIJI in order to determine neurite density (neurite area/explant area). Validation information for commercial antibodies is available on the manufacturers' websites. All experiments were performed in groups of five independent cultures and repeated at least twice.

Cell culture

HEK293 EBNA cells as well as the mouse melanoma B16-F1 cell line were obtained from ATCC. Human dermal microvascular endothelial cells (HDMEC) and human lymphatic endothelial cells (HDLEC) were obtained from Promocell, human umbilical vain endothelial cells (HUVEC) from Lonza. The endothelial cells were cultured in Vasculife VEGF-Mv Medium Complete Kit (Cell Systems) according to the manufacturer´s instructions. The mouse melanoma cell line (HCMel12) was kindly provided by T. Tüting (University of Magdeburg)[62]. The mouse melanoma cell lines HCMel12 and B16-F1 as well as HEK293 EBNA cells were cultured in Dulbecco´s Modified Eagle Medium (Gibco) supplemented with 10% fetal bovine serum (FBS). Pericyte-like vascular cells (PVCs) were isolated from meninigi of the Anxa5-LacZ reporter mouse strain[63]. PVCs were cultured in DMEM supplemented with 10% FCS and used at passages 28-40.

Tube-like formation assay

Cells were seeded at a density of 25,000 cells per well onto chamber slides (Labtek) or ibidi slides (ibidi) previously coated at 37°C with Matrigel™ BM matrix (Corning, New York, United States) for 30 min to allow polymerization. For the tube formation assay, the indicated proteins were added directly after seeding the cells. Cells were incubated for 14 h. For the tube regression assay proteins were added 7 h after seeding the endothelial cells, which had established a tubular network. Recombinant proteins were added at the concentrations indicated and in terms of endothelial cells treatment PDGF as well as VEGF (Biomol) was added at final concentrations of 10 ng ml^{-1}. Tube formation was documented by bright-field microscopy using a Nikon Eclipse TE2000-U microscope. In order to investigate the time course of tube formation, we performed time-lapse microscopy using an Olympus IX81. All experiments were performed in groups of five independent cultures and repeated at least three times.

Three-dimensional culture in collagen I matrix

Endothelial and pericyte-like cells (HUVECs/PVCs) were co-cultured in 2% collagen I gel as described previously[64]. Cultures were incubated in ECGM2 medium supplemented with PDGF-BB and VEGF (10 ng ml^{-1} each) for 6 days. Recombinant Net4-FL and Net4-FL^{E195AR199A} were added to the 3D gel solution at indicated concentration. Medium (supplemented with PDGF-BB, VEGF, and Net4-FL or Net4-FL^{E195AR199A}) was replaced every 48 h. Samples were fixed with Dent’s fixative and immunostained for collagen IV (AB769, Chemicon), laminin γ1 (kind gift from U. Mayer, University of East Anglia), and human CD31 (555444, BD Biosciences)[65]. DNA was stained by Hoechst 33258. Overview pictures were taken from each sample by using a Zeiss Axioplan microscope and average tube length was calculated by the Volocity software. Optical section images were taken by a Zeiss Apotome microscope and processed by Axiovision program. All experiments were performed in groups of five independent cultures and repeated at least three times. Validation information for the laminin γ1 antibody is presented in the work of Mayer et., al.[66].

Chick Chorioallantoic Membrane angiogenesis assay

The *ex ovo* culture of the chicken embryos was adapted from Auerbach et al.[46]. Briefly, fertilized Ross chicken eggs (Wüthrich Brüterei AG, Switzerland) were incubated at 37°C for 3 days in humidified atmosphere. On day 3, the eggs were carefully opened and their content was transferred into 100 x 20 mm Petri dishes (Corning, Switzerland). The chicken embryos were incubated in the same conditions for 6 more days. On day 9, the PEG hydrogels containing no protein and different proteins (Net4-FL, Net4-FL^{E195A,R199A}) were placed on the surface of the Chick Chorioallantoic Membrane (CAM). 15 µL PEG hydrogels were formed by Factor XIII$_a$-catalyzed cross-linking of two 8-arm PEG components (8-PEG–Gln and 8-PEG–MMP$_{sensitive}$–Lys) in 50 mM Tris pH 7.6, 50 mM $CaCl_2$ buffer to a final PEG concentration of 2% (w/v). Each hydrogel contained 4.25 µg of recombinant Net4-FL, or Net4-FL^{E195A,R199A}, which were added to the hydrogel mixture before polymerization. After 48 h of treatment, 100 µl of 2.5% FITC-dextran (MW 20,000, Sigma-Aldrich, Switzerland) in 0.9% NaCl solution were injected intravenously in the CAM. The vasculature was visualized by fluorescence microscopy and images around the treatment site were collected. All experiments were performed in groups of five independent cultures and repeated at least twice.

The images collected by fluorescence microscopy were analyzed using the ImageJ software (ImageJ 1.48, http://imagej.nih.gov/) using a script based on the quantification method developed by Blacher et al.[67]. The background of each image was evened out and the contrast enhanced. Then the threshold was adjusted manually to discriminate the vascularized area from the non-vascularized area, and the image was transformed into a binary image. A series of filters were applied then to refine the capillary structures. ROIs were set manually to exclude second-order and higher-ranking blood vessels and regions with underlying blood vessels. The vascularized areas within the ROIs, as well as the total ROI area, were measured. The obtained values were used to determine the distribution vascularized areas in the capillary network. All experiments were performed in groups of five independent cultures and repeated at least three times.

Immunohistochemistry of the CAM

After live imaging, the CAMs were fixed with 4% para-formaldehyde solution by applying the solution both on top and beneath the CAM membranes and incubated at room temperature for 1 h. The area around the treatment site was carefully cut out and placed between cellulose sheets before being processed for paraffin embedding. Tissue sections of 4 µm were cut with a microtome and mounted on glass slides for hematoxylin and eosin staining.

Transmission electron microscopy analyses of CAM capillaries

After live imaging, CAMs were fixed with 2.5% glutaraldehyde (buffered with 0.1M sodium-cacodylate, pH 7.4, 540 mOsm.) by applying the solution both on top and beneath the CAM membranes. They were post-fixed in 1% OsO_4 (0.1M sodium-cacodylate, pH 7.4, 340 mOsm), dehydrated in ethanol, and embedded in Epon 812 (Fluka, Buchs, Switzerland). Semi-thin sections were prepared, stained with toluidine blue, and analyzed with a light microscope[68,69]. Representative areas were selected and further imaged by transmission electron microscopy. Tissue samples from groups of three of three independent biological replicates were analyzed.

Therapeutic treatment of melanoma

The melanoma HCmel12 cell line, which spontaneously develops lung metastases after intracutaneous injection in immunocompetent C57BL/6 mice, was established from a primary Hgf-Cdk4R24C melanoma[62]. For tumor transplantation 2×10^5 cells in 100 µl 1 x PBS were injected intracutaneously into the flanks of C57BL/6 mice (female, $n = 5$, 8 weeks of age) and tumor development was monitored by visual inspection and palpation. When tumors exceed 2 mm groups of five mice were peritumorally injected daily with 1 µM of mouse albumin (ctrl), Net4-ΔC, Net4-FL, and Net4-FL^{E195A,R199A} for ten consecutive days. Tumor sizes were measured every second day and recorded as mean diameter in mm. Mice with tumors exceeding 20 mm diameter were sacrificed. All experiments were performed in groups of five mice and repeated independently at least twice.

Statistics

Statistical analysis was performed using the GraphPad Prism software (GraphPad Prism 5.00). The two sided *t*-test was used for statistical analysis, statistical analyses of the Schwann cell assay was performed using the one-way Anova followed by a pairwise comparison using the Holm-Sidak method. Results are displayed as mean ± s.d. except the one of the neurite outgrowth assay, in which the results are displayed as ± s.e.m..

References

1. Ishii, N., Wadsworth, W.G., Stern, B.D., Culotti, J.G. & Hedgecock, E.M. UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in C. elegans. *Neuron* **9**, 873-81 (1992).

2. Kennedy, T.E., Serafini, T., de la Torre, J.R. & Tessier-Lavigne, M. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425-35 (1994).

3. Koch, M. et al. A novel member of the netrin family, beta-netrin, shares homology with the beta chain of laminin: identification, expression, and functional characterization. *J Cell Biol* **151**, 221-34 (2000).

4. Nakashiba, T. et al. Netrin-G1: a novel glycosyl phosphatidylinositol-linked mammalian netrin that is functionally divergent from classical netrins. *J Neurosci* **20**, 6540-50 (2000).

5. Nakashiba, T., Nishimura, S., Ikeda, T. & Itohara, S. Complementary expression and neurite outgrowth activity of netrin-G subfamily members. *Mech Dev* **111**, 47-60 (2002).

6. Yin, Y., Sanes, J.R. & Miner, J.H. Identification and expression of mouse netrin-4. *Mech Dev* **96**, 115-9 (2000).

7. Yamagishi, S. et al. Netrin-5 is highly expressed in neurogenic regions of the adult brain. *Front Cell Neurosci* **9**, 146 (2015).

8. Garrett, A.M. et al. Analysis of Expression Pattern and Genetic Deletion of Netrin5 in the Developing Mouse. *Front Mol Neurosci* **9**, 3 (2016).

9. Serafini, T. et al. The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. *Cell* **78**, 409-24 (1994).

10. Carmeliet, P. & Tessier-Lavigne, M. Common mechanisms of nerve and blood vessel wiring. *Nature* **436**, 193-200 (2005).

11. Keino-Masu, K. et al. Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell* **87**, 175-85 (1996).

12. Fazeli, A. et al. Phenotype of mice lacking functional Deleted in colorectal cancer (Dcc) gene. *Nature* **386**, 796-804 (1997).

13. Hong, K. et al. A ligand-gated association between cytoplasmic domains of UNC5 and DCC family receptors converts netrin-induced growth cone attraction to repulsion. *Cell* **97**, 927-41 (1999).

14. Corset, V. et al. Netrin-1-mediated axon outgrowth and cAMP production requires interaction with adenosine A2b receptor. *Nature* **407**, 747-50 (2000).

15. Hopker, V.H., Shewan, D., Tessier-Lavigne, M., Poo, M. & Holt, C. Growth-cone attraction to netrin-1 is converted to repulsion by laminin-1. *Nature* **401**, 69-73 (1999).

16. Yebra, M. et al. Recognition of the neural chemoattractant Netrin-1 by integrins alpha6beta4 and alpha3beta1 regulates epithelial cell adhesion and migration. *Dev Cell* **5**, 695-707 (2003).

17. Ly, A. et al. DSCAM is a netrin receptor that collaborates with DCC in mediating turning responses to netrin-1. *Cell* **133**, 1241-54 (2008).

18. Eveno, C. et al. Netrin-4 delays colorectal cancer carcinomatosis by inhibiting tumor angiogenesis. *Am J Pathol* **178**, 1861-9 (2011).

19. Lejmi, E. et al. Netrin-4 inhibits angiogenesis via binding to neogenin and recruitment of Unc5B. *Proc Natl Acad Sci U S A* **105**, 12491-6 (2008).

20. Nacht, M. et al. Netrin-4 regulates angiogenic responses and tumor cell growth. *Exp Cell Res* **315**, 784-94 (2009).

21. Larrieu-Lahargue, F., Welm, A.L., Thomas, K.R. & Li, D.Y. Netrin-4 induces lymphangiogenesis in vivo. *Blood* **115**, 5418-26 (2010).

22. Finci, L.I. et al. The crystal structure of netrin-1 in complex with DCC reveals the bifunctionality of netrin-1 as a guidance cue. *Neuron* **83**, 839-49 (2014).

23. Xu, K. et al. Neural migration. Structures of netrin-1 bound to two receptors provide insight into its axon guidance mechanism. *Science* **344**, 1275-9 (2014).

24. Grandin, M. et al. Structural Decoding of the Netrin-1/UNC5 Interaction and its Therapeutical Implications in Cancers. *Cancer Cell* **29**, 173-185 (2016).

25. Schneiders, F.I. et al. Binding of netrin-4 to laminin short arms regulates basement membrane assembly. *J Biol Chem* **282**, 23750-8 (2007).

26. Staquicini, F.I. et al. Discovery of a functional protein complex of netrin-4, laminin gamma1 chain, and integrin alpha6beta1 in mouse neural stem cells. *Proc Natl Acad Sci U S A* **106**, 2903-8 (2009).

27. Timpl, R. & Brown, J.C. Supramolecular assembly of basement membranes. *Bioessays* **18**, 123-32 (1996).

28. Stratman, A.N., Malotte, K.M., Mahan, R.D., Davis, M.J. & Davis, G.E. Pericyte recruitment during vasculogenic tube assembly stimulates endothelial basement membrane matrix formation. *Blood* **114**, 5091-101 (2009).

29. Gerhardt, H. & Betsholtz, C. Endothelial-pericyte interactions in angiogenesis. *Cell Tissue Res* **314**, 15-23 (2003).

30. Poschl, E. et al. Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development* **131**, 1619-28 (2004).

31. Thyboll, J. et al. Deletion of the laminin alpha4 chain leads to impaired microvessel maturation. *Mol Cell Biol* **22**, 1194-202 (2002).

32. Miner, J.H., Cunningham, J. & Sanes, J.R. Roles for laminin in embryogenesis: exencephaly, syndactyly, and placentopathy in mice lacking the laminin alpha5 chain. *J Cell Biol* **143**, 1713-23 (1998).

33. Hussain, S.A., Carafoli, F. & Hohenester, E. Determinants of laminin polymerization revealed by the structure of the alpha5 chain amino-terminal region. *EMBO Rep* **12**, 276-82 (2011).

34. McKee, K.K., Harrison, D., Capizzi, S. & Yurchenco, P.D. Role of laminin terminal globular domains in basement membrane assembly. *J Biol Chem* **282**, 21437-47 (2007).

35. Purvis, A. & Hohenester, E. Laminin network formation studied by reconstitution of ternary nodes in solution. *J Biol Chem* **287**, 44270-7 (2012).

36. Stetefeld, J., Mayer, U., Timpl, R. & Huber, R. Crystal structure of three consecutive laminin-type epidermal growth factor-like (LE) modules of laminin gamma1 chain harboring the nidogen binding site. *J Mol Biol* **257**, 644-57 (1996).

37. Patel, T.R. et al. Determination of a molecular shape for netrin-4 from hydrodynamic and small angle X-ray scattering measurements. *Matrix Biol* **31**, 135-40 (2012).

38. Qin, S., Yu, L., Gao, Y., Zhou, R. & Zhang, C. Characterization of the receptors for axon guidance factor netrin-4 and identification of the binding domains. *Mol Cell Neurosci* **34**, 243-50 (2007).

39. Zhang, C. et al. Identification of a novel alternative splicing form of human netrin-4 and analyzing the expression patterns in adult rat brain. *Brain Res Mol Brain Res* **130**, 68-80 (2004).

40. Lange, J. et al. The axon guidance molecule Netrin-4 is expressed by Muller cells and contributes to angiogenesis in the retina. *Glia* **60**, 1567-78 (2012).

41. Wilson, B.D. et al. Netrins promote developmental and therapeutic angiogenesis. *Science* **313**, 640-4 (2006).

42. Dakouane-Giudicelli, M., Alfaidy, N. & de Mazancourt, P. Netrins and their roles in placental angiogenesis. *Biomed Res Int* **2014**, 901941 (2014).

43. Han, Y. et al. Therapeutic effects of topical netrin-4 inhibits corneal neovascularization in alkali-burn rats. *PLoS One* **10**, e0122951 (2015).

44. Lambert, E., Coissieux, M.M., Laudet, V. & Mehlen, P. Netrin-4 acts as a pro-angiogenic factor during zebrafish development. *J Biol Chem* **287**, 3987-99 (2012).

45. Francescone, R.A., 3rd, Faibish, M. & Shao, R. A Matrigel-based tube formation assay to assess the vasculogenic activity of tumor cells. *J Vis Exp* (2011).

46. Auerbach, R., Kubai, L., Knighton, D. & Folkman, J. A simple procedure for the long-term cultivation of chicken embryos. *Dev Biol* **41**, 391-4 (1974).

47. Liu, Y. et al. Novel role for Netrins in regulating epithelial behavior during lung branching morphogenesis. *Curr Biol* **14**, 897-905 (2004).

48. Han, Y. et al. Therapeutic effects of topical netrin-4 in a corneal acute inflammatory model. *Int J Ophthalmol* **8**, 228-33 (2015).

49. Jain, R.K. Molecular regulation of vessel maturation. *Nat Med* **9**, 685-93 (2003).

50. Suri, C. et al. Requisite role of angiopoietin-1, a ligand for the TIE2 receptor, during embryonic angiogenesis. *Cell* **87**, 1171-80 (1996).

51. Yoder, E.J. Modifications in astrocyte morphology and calcium signaling induced by a brain capillary endothelial cell line. *Glia* **38**, 137-45 (2002).

52. Esseghir, S. et al. Identification of NTN4, TRA1, and STC2 as prognostic markers in breast cancer in a screen for signal sequence encoding proteins. *Clin Cancer Res* **13**, 3164-73 (2007).

53. Gasteiger E. et al. Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* (ed. Walker, J.M.) 571-607 (Humana Press, 2005).

54. Artimo, P. et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* **40**, W597-603 (2012).

55. Carafoli, F., Hussain, S.A. & Hohenester, E. Crystal structures of the network-forming short-arm tips of the laminin beta1 and gamma1 chains. *PLoS One* **7**, e42473 (2012).

56. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

57. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).

58. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).

59. Murshudov, G.N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-67 (2011).

60. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).

61. Wienken, C.J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat Commun* **1**, 100 (2010).

62. Bald, T. et al. Ultraviolet-radiation-induced inflammation promotes angiotropism and metastasis in melanoma. *Nature* **507**, 109-13 (2014).

63. Brachvogel, B. et al. Perivascular cells expressing annexin A5 define a novel mesenchymal stem cell-like population with the capacity to differentiate into multiple mesenchymal lineages. *Development* **132**, 2657-68 (2005).

64. Cooley, L.S. et al. Reversible transdifferentiation of blood vascular endothelial cells to a lymphatic-like phenotype in vitro. *J Cell Sci* **123**, 3808-16 (2010).

65. Bader, B.L. et al. Compound genetic ablation of nidogen 1 and 2 causes basement membrane defects and perinatal lethality in mice. *Mol Cell Biol* **25**, 6846-56 (2005).

66. Mayer, U. et al. A single EGF-like motif of laminin is responsible for high affinity nidogen binding. *EMBO J* **12**, 1879-85 (1993).

67. Blacher, S. et al. Quantification of in vivo tumor invasion and vascularization by computerized image analysis. *Microvasc Res* **75**, 169-78 (2008).

68. Djonov, V.G., Galli, A.B. & Burri, P.H. Intussusceptive arborization contributes to vascular tree formation in the chick chorio-allantoic membrane. *Anat Embryol (Berl)* **202**, 347-57 (2000).

69. Matschke, K., Da Silva-Azevedo, L., Hlushchuk, R., Djonov, V. & Baum, O. Annexins as cell-type-specific markers in the developing chicken chorionallantoic membrane. *Cell Tissue Res* **323**, 395-404 (2006).

Chapter 4: Dramatic and concerted conformational changes enable rhodocetin to block α2β1 integrin selectively

Johannes A. Eble[1*#], Matthew McDougall[2#], George L. Orriss[2], Stephan Niland[1], Benjamin Johanningmeier[1], Gottfried Pohlentz[3], Markus Meier[2], Simone Karrasch[2], Maria-Inacia Estevão-Costa[1], Augusto Martins Lima[1], Jörg Stetefeld[2*]

# These two authors contributed equally.

[1]Institute of Physiological Chemistry and Pathobiochemistry, University of Münster,
Waldeyerstraße 15, 48149 Münster, Germany

[2]Departments of Chemistry & Microbiology, University of Manitoba, 144 Dysart Road, R3T 2N2 Winnipeg, MB, Canada

[3]Institute of Hygiene, University of Münster, Robert-Koch-Straße 41, 48149 Münster, Germany

First published in *Plos Biology*, **7,** Article e2001492 (2017)

DOI 10.1371/journal.pbio.2001492

Preface

Although X-ray diffraction yields static images of a crystallized protein, it can give insight into the structural reorganizations that accompany macromolecular function. In the original rhodocetin structure, the 4 subunits are in a cruciform arrangement, but the αβ and γδ heterodimers have distinct targets and must dissociate for these interactions to occur. The crystal structure of the γδ heterodimer of rhodocetin in complex with the vWF A domain of α2 integrin reveals that the snake venom protein undergoes dramatic conformational changes to bind its target and block the blood clotting response.

Author Contributions

Rhodocetin and its complexed were isolated and purified by Johannes A. Eble, who also performed the binding assays and titration curves. Simone Karrasch crystallized the RCγδ-α2A complex, and George L. Orriss collected the diffraction data. Matthew McDougall with Jörg

Stetefeld and Markus Meier evaluated the structural data. Stephan Niland and Benjamin Johanningmeier carried out the molecular biological cloning steps and set up the expression and purification of the α2A mutants. Augusto Martins Lima, along with Maria-Inacia Estevão-Costa, helped with the purification of α2A mutants and with the chemical modification of RC. Gottfried Pohlentz indentified the epitope of IIIG5 by mass spectrometry. The manuscript was prepared by Matthew McDougall, George L. Orriss, Jörg Stetefeld, and Johannes A. Eble.

Abstract

The collagen binding integrin α2β1 plays a crucial role in hemostasis, fibrosis, and cancer progression amongst others. It is specifically inhibited by rhodocetin (RC), a C-type lectin-related protein (CLRP) from the Malayan pit viper (*Calloselasma rhodostoma*) venom. The structure of RC alone reveals a heterotetramer arranged as an αβ and γδ subunit in a cruciform shape. RC specifically binds to the collagen binding A-domain of the integrin α2 subunit, thereby blocking collagen induced platelet aggregation. However, until now the molecular basis for this interaction has remained unclear. Here we present the molecular structure of the RCγδ-α2A complex solved to 3.0 Å resolution. Our findings show that rhodocetin undergoes a dramatic quaternary reassembly upon binding to α2β1 integrin. Besides the release of the non-binding RCαβ tandem, the RCγ subunit interacts with loop 2 of the α2A domain as result of a dramatic conformational change. The RCδ subunit contacts the integrin α2A domain in the "closed" conformation through its helix C. Combined with epitope-mapped antibodies,

conformationally locked α2A domain mutants, point mutations within the α2A loop 2, and chemical modifications of the purified toxin protein, this molecular structure of RCγδ-α2A complex explains the inhibitory mechanism and specificity of RC for α2β1 integrin.

Introduction

Most cellular processes depend on the formation of interactions between cells and the extracellular matrix (ECM). Key facilitators of these interactions are the integrins. They consist of two subunits, α and β, each of which has multiple isoforms[1,2]. The different subunit composition between integrins determines their ligand binding specificity and functionality. Integrins are cell adhesion molecules, which are involved in a broad range of cell functions, such as proliferation, differentiation, adhesion and migration. Defect or dysfunction of integrins, in particular of α2β1 integrin, a prominent collagen binding receptor of many cell types[3] and the only collagen binding integrin on platelets[4] may affect vascular development and angiogenesis[5], epithelial cell differentiation[6], wound repair and fibrosis[7], inflammation[8,9], cancer and cancer therapy[10], as well as collagen-induced platelet activation, hemostasis and thrombosis[4,11]. Therefore, α2β1 integrin has become a prominent target in drug research[12-14].

The collagen binding site is located within the α2A domain of α2β1 integrin which is homologous to the A-domain of von Willebrand factor (vWF). The α2A domain contains a metal ion that is required for collagen binding as it part of the binding site for the collagen triple helix[15]. In order to bind to collagen, the α2A domain undergoes a series of concerted conformational changes. In short, helix C unwinds, the N-termini of helices 6 and 7 simultaneously turn away from each other and finally helix 7 moves downward against helix 1 to give the collagen binding “open” conformation which contrasts with the previous “closed” conformation[15, 16]. This likely general mechanism of molecular movement of integrin A-domains was subsequently confirmed by introducing a disulfide bridge into the A-domain of the integrin αL subunit such that this interconversion was blocked with the protein locked in either the “open” or “closed” state[17].

Integrin function can be blocked by two major classes of snake venom proteins, the disintegrins[18,19] and the C-type lectin-related proteins (CLRPs)[20,21]. In contrast to the disintegrins which can target multiple integrins, CLRPs specifically inhibit α2β1 integrin activity[21]. The high selectivity and affinity of these snake venom proteins for α2β1 integrin makes them ideal lead compounds for drug development[22-24]. Current members of the CLRP family include the proteins rhodocetin, EMS16, vixapatin, sochicetin-B, lebecetin, flavocetin and rhinocetin[25-31]. As more CLRP protein structures become available it is clear that although the supramolecular structure can vary from the basic heterodimer of EMS16[27] to the ring like $(\alpha\beta)_4$ structures of flavocetin and convulxin[32,33] the underlying basic unit is a heterodimer consisting of two subunits, usually named α and β, which dimerize via their characteristic index finger loops[20,34]. Interestingly, in the case of the rhodocetin heterotetramer (αβγδ) structure[26], the αβ and γδ subunits form two heterodimeric pairs that are orientated orthogonally towards each other in a cruciform shape. Despite these differences, the subunits of CLRP family members are highly homologous with each other. Evolutionarily, the CLRP fold has developed from a carbohydrate recognizing domain (CRD) into a structure which specifically targets clotting factors IX and X, α2β1 integrin, and other platelet adhesion receptors[20,34-36]. Among the latter, the vWF receptor and the two collagen binding receptors, glycoprotein GPIV and α2β1 integrin, are targets for snake venom CLRPs, thereby inhibiting or activating platelet activation and aggregation[37,38]. Consequently, these snake venom proteins severely interfere with hemostasis[36,39]. However, the nature of the molecular mechanism by which CLRPs inhibit α2β1 integrin and by which CLRPs implement specificity towards α2β1 integrin has remained undetermined.

Rhodocetin (RC) is a CLRP of the Malayan pit viper *Calloselasma rhodostoma*[26] and, together with EMS16 from *Echis multisquamatus*, they are the only known CLRP family members proven to target the α2A domain for which atomic resolution structures are available[27,40]. Unlike the α2β1 integrin-collagen interaction, which is metal ion dependent, the binding of RC to α2β1 integrin does not require a metal ion, which implies a different mechanism of action. In a previous study, we demonstrated that the RCαβγδ heterotetramer binds to α2β1 integrin before releasing the αβ subunit (RCαβ) from the complex[40]. In the current work, we present the molecular structure of this RCγδ-α2A domain complex and unravel the molecular mechanism of this interaction. The RC binding site overlaps with that of collagen including the key metal ion site, thereby sterically blocking collagen binding. Moreover, a comparison with the previously determined rhodocetin structure[26] reveals that, in addition to the release of the RCαβ subunit, the RCγδ subunit undergoes a major conformational change upon integrin binding which causes it to snap into a bent conformation like a mouse trap. In this final state, RCγδ holds the α2A domain in the “closed” conformation, allosterically unable to bind to collagen. The result is a highly efficient inhibition of α2β1 integrin mediated attachment and signaling in cells and platelets.

Results

Purification and characterization of the RCγδ-α2A complex

To isolate rhodocetin (RC) in complex with the integrin α2A domain, recombinant α2A domain was immobilized to Ni Sepharose resin via its His_6-tag. Thereafter, a RC-rich protein fraction of *Calloselasma rhodostoma* venom was applied to this column, resulting in the formation of the complex of α2A with tetrameric RC (RCαβγδ) which still bound to the column. Treatment with 5 mM EGTA resulted in the dissociation of the α2A domain bound RC tetramer and the release of RCαβ from the complex, which was eluted from the column. In contrast, RCγδ remained firmly attached to the column bound α2A (**Figure 4-1**). This RCγδ-α2A complex was then eluted with a linear gradient of imidazole (**Figure 4-1a**). Its His_6-tag was cleaved by trypsinolysis and the excess α2A was removed by size exclusion chromatography. The close physical contact of both partners within the RCγδ-α2A complex was proven by cross-linkage with 0.5 mM bis(sulfosuccinimidyl)suberate (BS^3) (**Figure 4-1b**).

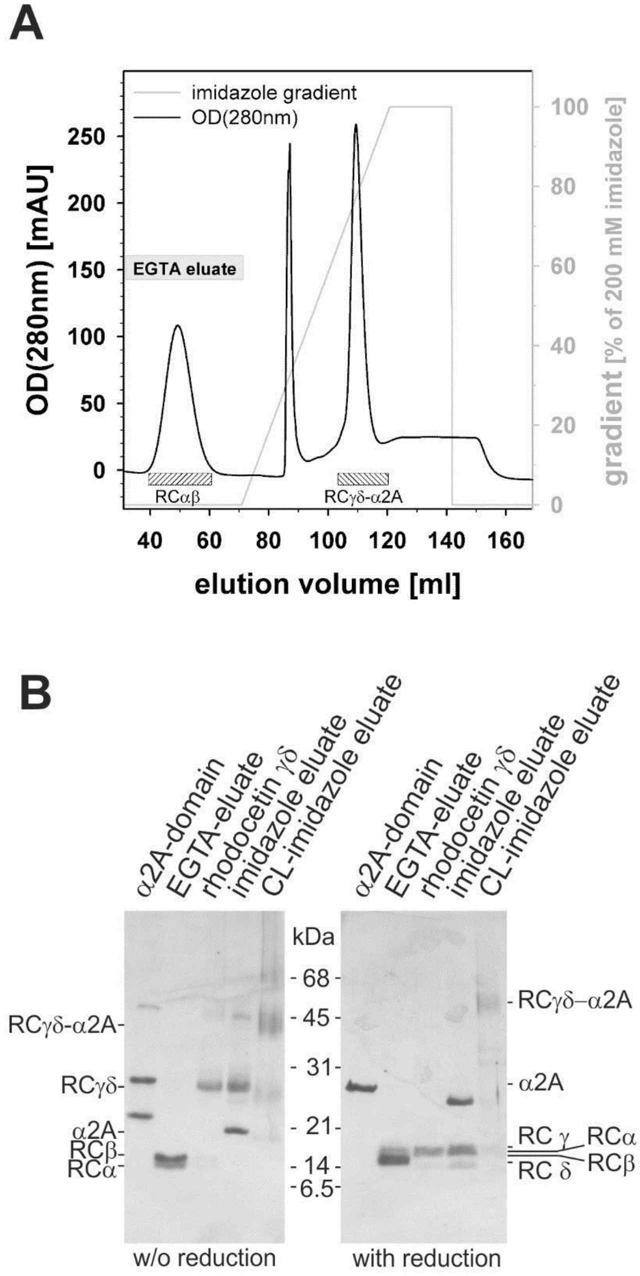


Figure 4-1. Isolation of the rhodocetin γδ-α2A complex on Ni Sepharose column.

(**a**) Elution profile of the Ni Sepharose affinity chromatography column. The RCγδ-α2A complex was formed on a Ni Sepharose column by subsequently loading the oligo His-tagged α2A domain and RCαβγδ. RCαβ and the RCγδ-α2A complex were eluted with EGTA and an imidazole gradient, respectively. (**b**) SDS-PAGE of eluate fractions (lanes "EGTA eluate" and "imidazole eluate"), in comparison to isolated control proteins (lanes "α2A domain" and "rhodocetin γδ"), under non-reducing and reducing conditions and stained with silver. Note that the trypsin-trimmed RCγδ-α2A complex showed a slightly reduced size of the α2A domain due to the proteolytic removal of the His_6-tag. The physical contact of co-eluted rhodocetin γδ and α2A domain was analytically proven by cross-linkage with 0.5 mM BS^3 (lane "CL-imidazole eluate").

Molecular structure of the rhodocetin γδ-α2A complex

The crystal structure of the RCγδ-α2A complex was determined at 3.0 Å resolution by molecular replacement using the previously determined RCαβγδ structure (pdb:3GPR) as a search template (**Figure 4-2**). The RCγδ-α2A structure clearly showed that the RCγδ subunit bound to the top of the α2A domain directly above the metal ion-binding site thereby sterically blocking access of collagen (**Figure 4-2a**). Both chains of RCγδ are typical CLRP folds, characterized by a globular core domain inter-linked mutually by extended index finger loops. The A-domain of α2β1 integrin assumed the "closed" conformation with its central β-sheet flanked by the α-helices 3, 1, 7 and 4, 5, 6 on either side. The crystal structures contains six RCγδ-α2A complexes per asymmetric unit (**Appendix 2, Figure II-1**)

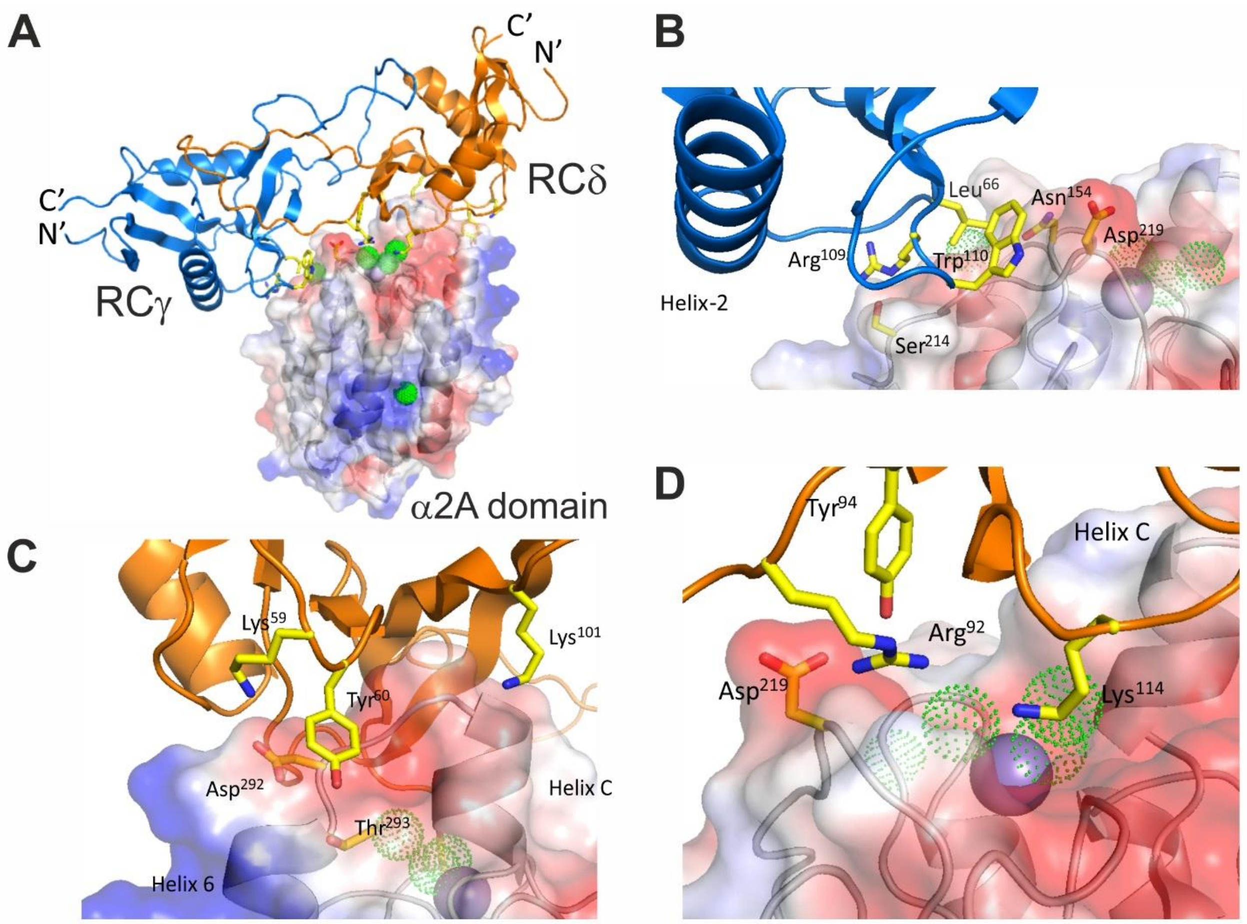


Figure 4-2. The molecular structure of the RCγδ-α2A complex.

(**a**) Molecular structure of the RCγδ-α2A complex reveals that RCγδ binds on the top and lateral face of the α2A domain. The RCγδ subunit covers the collagen binding crevice of the α2A domain, with its long axis perpendicular to the collagen ligand interaction. (**b**) Detailed view of the interaction site between the RCγ chain and loop 2 of α2A. (**c,d**) Two different views of the interaction site between the RCδ subunit and helix C of α2A. The α2A domain is shown as a transparent surface in (**a**) thru (**d**), with the key binding residues labelled.

We determined the total interaction surface between RCγδ and α2A in the complex to be 965 Å^2.

There were two interface areas on the surface of RCγδ in contact with α2A (**Figure 4-2b-d**).

First, the larger interaction site (715 Å^2) consisted of two adjacent patches of three residues each on the RCδ subunit, K59-Y60-K101 (**Figure 4-2c**), and R92-Y94-K114 (**Figure 4-2d**), which were largely hydrophilic. Second, a smaller hydrophobic site (280 Å^2) on the RCγ subunit consisted of the triad L66-R109-W110 that interacted with helix 3, helix 4 and loop 2 of α2A (**Figure 4-2b**).

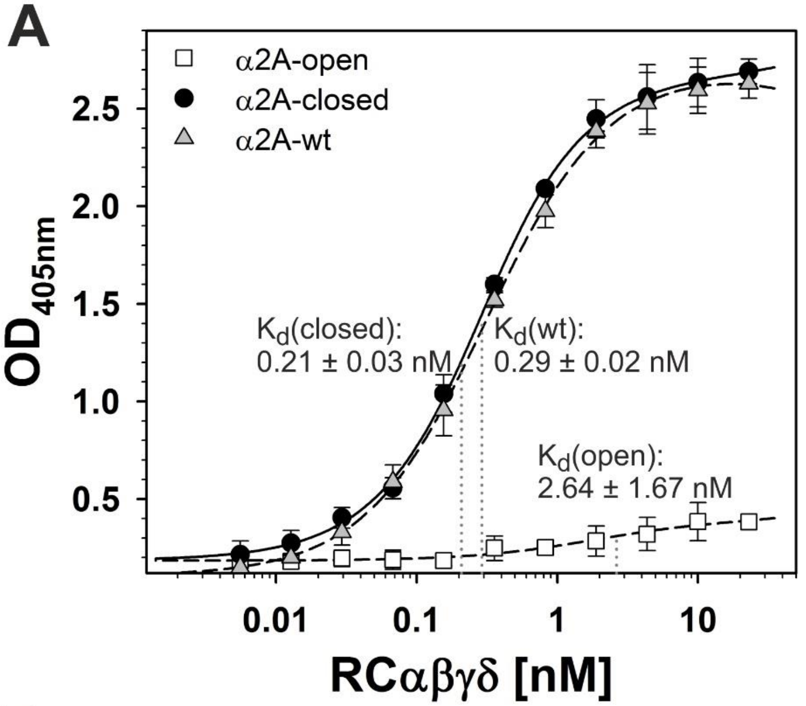
Two complementary contact surfaces on the α2A domain extended down from helix C and the metal ion-binding site (top face) to the loop 2 sequence S^{214}QYGGD^{219} (lateral face) to form an almost contiguous interface which interacted with the RCγδ subunit. The top face of α2A was approached by the RCδ subunit with its larger two patches containing interface (**Figure 4-2c, d**). The first patch comprised residues K59, Y60, and K101 of RCδ interacting with residues D292 and T293 together with the adjacent helix C of α2A. The side chains of K59 and Y60 were countered by complementary carboxylate and hydroxyl groups of D292 and T293 of α2A, while the amino group of K101 pointed towards the backbone carbonyl groups at the C-terminus of helix C. The second patch had the side chains of R92, Y94, and K114 of the RCδ subunit pointing into the collagen binding crevice of α2A. The long side chain of K114 of this protuberance sat at the entrance to the divalent cation binding site (**Figure 4-2d**) and was positioned 7.7 Å above the magnesium ion, whereas the positively charged guanidino group and the phenolic hydroxyl group of R92 and Y94 contacted the main chain carbonyl of D219 in the loop 2 of α2A.

The second contact surface is the loop 2 sequence $S^{214}QYGGD^{219}$ at the lateral face of α2A which interacted with the amino acid side chains of L66, R109, and W110 of the RCγ subunit (**Figure 4-2b**). For example, the aromatic indole ring of W110 contributed to a hydrophobic surface and interacted with the backbone chain of the glycine residues G217 and G218 together with the adjacent aspartate residue D219 within loop 2 of the α2A domain (**Figure 4-2b**). In addition, L66 of RCγ contacted N154 of loop 1 of the α2A domain. The final RCγ residue of the triad R109 made contacts to the S214 side chain of α2A. Taken together, the hydrophobic patch of the RCγ subunit predominantly interacted with the loop 2 sequence $S^{214}QYGGD^{219}$ of α2A. This loop 2 sequence immediately preceded residue T221 which was part of the metal ion binding site of α2A. A key residue with regard to the interface between the RCγδ subunit and the α2A domain in the RCγδ-α2A complex was the loop 2 D219 of α2A as it was part of both RC contact sites. In addition, it connected the loop 2 sequence with the collagen binding crevice and the helix C of α2A. The presence of helix C in the RCγδ-α2A complex structure indicated that RC had trapped the α2A domain in the "closed" conformation which is not capable of binding collagen[15].

Rhodocetin RCγδ binds the "closed" conformation of α2A

To test whether RC exclusively binds the closed conformation of α2A, we generated two conformationally distinct mutants, in which the A-domain was held by a disulfide bridge between K168C-E318C and K168C-A325C in the open and closed conformation, respectively

(**Appendix 2, Figure II-2**)[17,41]. Before introducing cysteine residues at these positions, it was necessary to replace the naturally occurring original cysteine residues at position 150 and 270 with alanines. No change in binding affinity to rhodocetin was observed for this α2A-C150A,C270A double mutant (data not shown). Into this cysteine free α2A domain, K168 of α-helix 1 was replaced by a cysteine residue with a second cysteine residue introduced into α-helix 7 at either position E318 or A325. As a consequence of the newly formed disulfide bridge, the movement of helices 1 & 7 with respect to each other that occurred when α2A shifts between the "open" and "closed" conformation was blocked. Thus, the α2A domain was held in the "open" (K168C-E318C) and "closed" (K168C-A325C) conformation, respectively. The α2A mutant with the "open" conformation hardly bound to RC (**Figure 4-3a**), while RC binding to the "closed" conformation of α2A (K_d-value: 0.21 ± 0.03 nM) was similar to that obtained with wild-type α2A (K_d-value: 0.29 ± 0.02 nM). Our structural findings revealed that the sidechain moiety of Lys101 is oriented towards the negatively charged dipole of helix C, stabilizing the closed conformation of the α2A domain (**Figure 4-3b**)



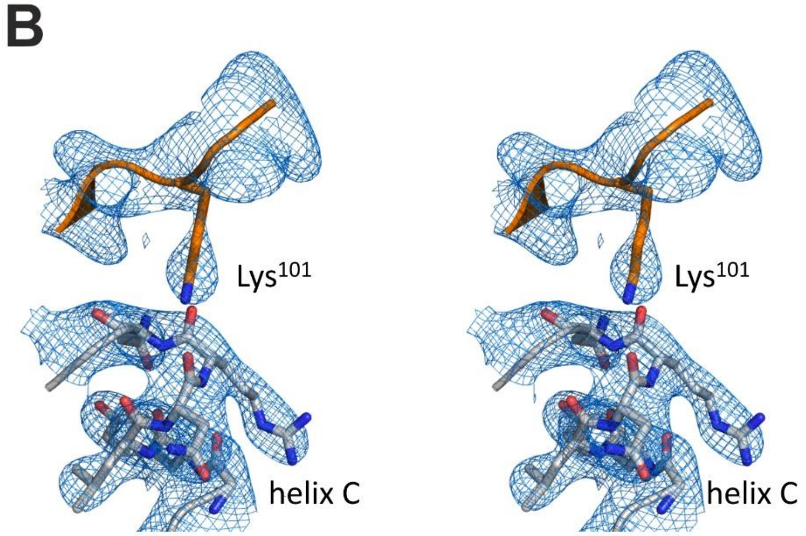


Figure 4-3. Rhodocetin recognizes the “closed” conformation, but not the “open” conformation of the integrin α2A domain.

(**a**) Titration of different α2A conformations with RC. The disulfide-locked conformation mutants, α2A open (☐) and α2A closed (●), were immobilized to microtiter plates at 10 µg/ml.

Along with immobilized α2A wild type form (△), they were titrated with RCαβγδ. Bound rhodocetin was fixed and quantified with a rabbit RC-antiserum by ELISA with a photometric signal at 405 nm. The OD_{405} values were corrected for α2A domain free, BSA blocked controls. The data presented here are taken from three independent experiments, with each measurement made in duplicate. Means ± S.D. (n=6) are shown. The K_d values for RC binding to the disulfide-locked conformation mutants and the wild type (wt) form of α2A are indicated at the titration curves. Both "open" and "closed" conformations have significantly different K_d values when compared to the one of the wild type form (indicated by *, $p<0.05$, Student´s t-test). (**b**) The crystal structure of the RCγδ-α2A complex reveals that the "closed" conformation of the α2A domain with its characteristic helix C is stabilized by the bound RCδ subunit. A stereo view of the Sigma-A weighted 2Fo-Fc map at 3.0Å resolution is shown at 1.5σ contour level.

The epitope of the monoclonal antibody IIIG5 is unmasked in the RCγδ-α2A complex

Among several monoclonal antibodies raised against the RCγδ subunit[40], IIIG5 belonged to the subgroup that only recognized its epitope within RCγδ after its complexation with α2A and the subsequent release of the RCαβ subunit (**Figure 4-4a**). This became evident when the antibody was immobilized and its ability to capture RCαβγδ, RCγδ-α2A, or RCγδ out from solution was probed. IIIG5 gave a binding signal with the RCγδ-α2A complex and RCγδ, but not with the RC tetramer alone. Of the two rhodocetin species capable of binding the IIIG5 antibody, the RCγδ subunit gave the highest binding signal (**Figure 4-4a**). The most probable explanation for these results was that the IIIG5 epitope was fully accessible in RCγδ and so we observed what approximates the maximal binding. At the other extreme, we had no binding of RCαβγδ as the epitope was entirely masked in the tetramer. Between these two extremes was the RCγδ-α2A complex, where the epitope is sufficiently exposed for IIIG5 to bind but not to the same extent as for RCγδ due to the nature of the RCγδ-α2A interaction.

The sequence epitope of IIIG5 was isolated from a tryptic digestion of RCαβγδ by affinity chromatography on a IIIG5 column and subsequently by reversed phase-HPLC. Mass spectrometry identified the γ chain sequence 94-106 as the IIIG5 epitope (**Appendix 2, Figure II-3**), which was mainly located within the index finger loop of RCγ (**Figure 4-4b**). This result can be clearly explained by comparing the native RCαβγδ structure with the newly determined RCγδ-α2A complex structure. The IIIG5 epitope was covered by the RCαβ subunit in the RCαβγδ structure and only became accessible upon formation of the RCγδ-α2A complex. Moreover, the index finger loop of the RCγ underwent a major conformational change upon formation of the RCγδ-α2A complex, leading to increased accessibility of the IIIG5 epitope.

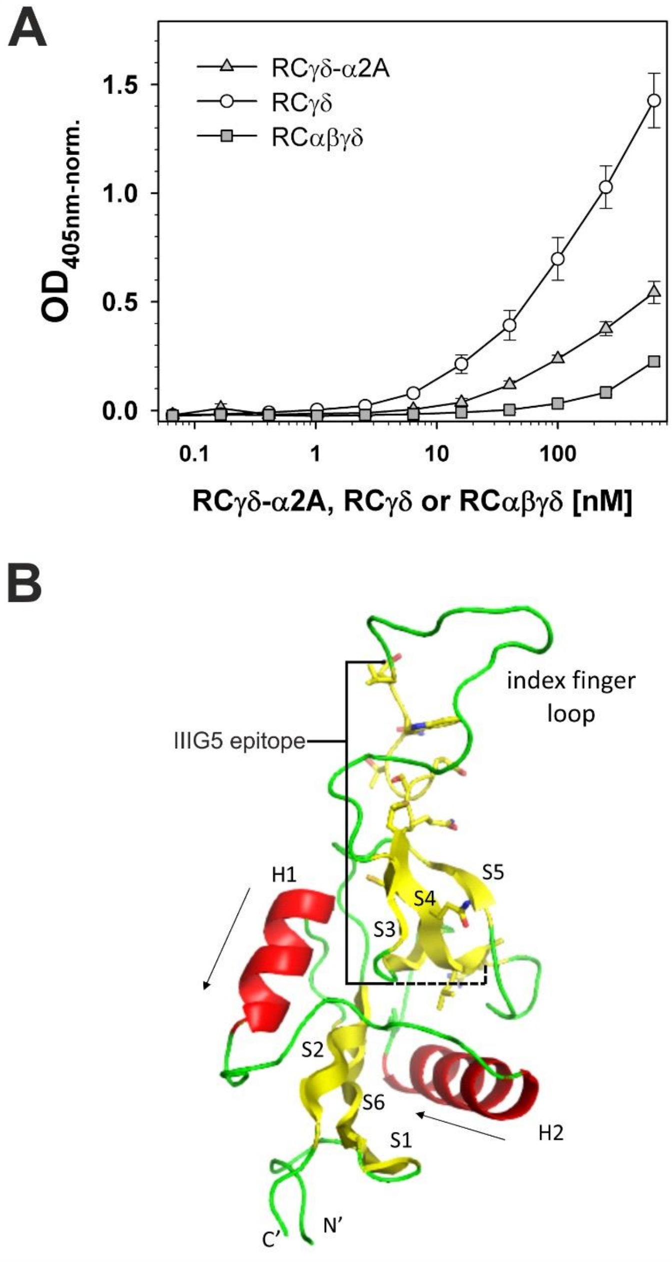


Figure 4-4.The monoclonal antibody IIIG5 recognizes its epitope within the RCγ subunit in the RCγδ-α2A complex but not in the tetrameric RCαβγδ.

(**a**) The monoclonal antibody IIIG5 recognized an epitope of the RCγ subunit, which is fully accessible in the RCγδ subunit (○), partially accessible in the RCγδ-α2A complex (light gray

▲) and completely covered in the RCαβγδ tetramer (dark gray ■). IIIG5 was immobilized onto microtiter plates and titrated with RCαβγδ, RCγδ-α2A complex or RCγδ subunit. Bound RC components were fixed and detected using rabbit RC antiserum with ELISA at 405 nm. The data presented here are taken from three independent experiments with each measurement done in duplicate. Means ± S.D. are shown. (**b**) Molecular structure of the CLRP-fold typical of all four rhodocetin chains. Both the γ and δ subunits of RC are very similar (Cα-RMSD 0.8Å) and feature a core structure with two α-helices (H1 and H2) flanked by two antiparallel β-sheets S1-S2-S6 and S3-S4-S5. The amino acid residues V94-R109 of the IIIG5 epitope of RCγ are highlighted.

Conformational changes within the RCγδ-dimer after α2A binding

The dramatic conformational changes that took place within the RCγδ subunit were readily apparent upon comparing the molecular structures of the RCγδ-α2A complex with the native RCαβγδ tetramer (**Figure 4-5**). The binding face of the RCαβγδ changes from a flat surface into concave binding surface to embrace the α2A domain (**Figure 4-5a, b**). This was implemented via (i) a rigid body movement of both core segments of chains γ and δ, (ii) a dramatic re-orientation of the index-finger loop of the γ-subunit, which harbors the IIIG5 epitope, and consequently (iii) local re-orientations of key binding residues in both RC subunits (**Figure 4-5c, d**).

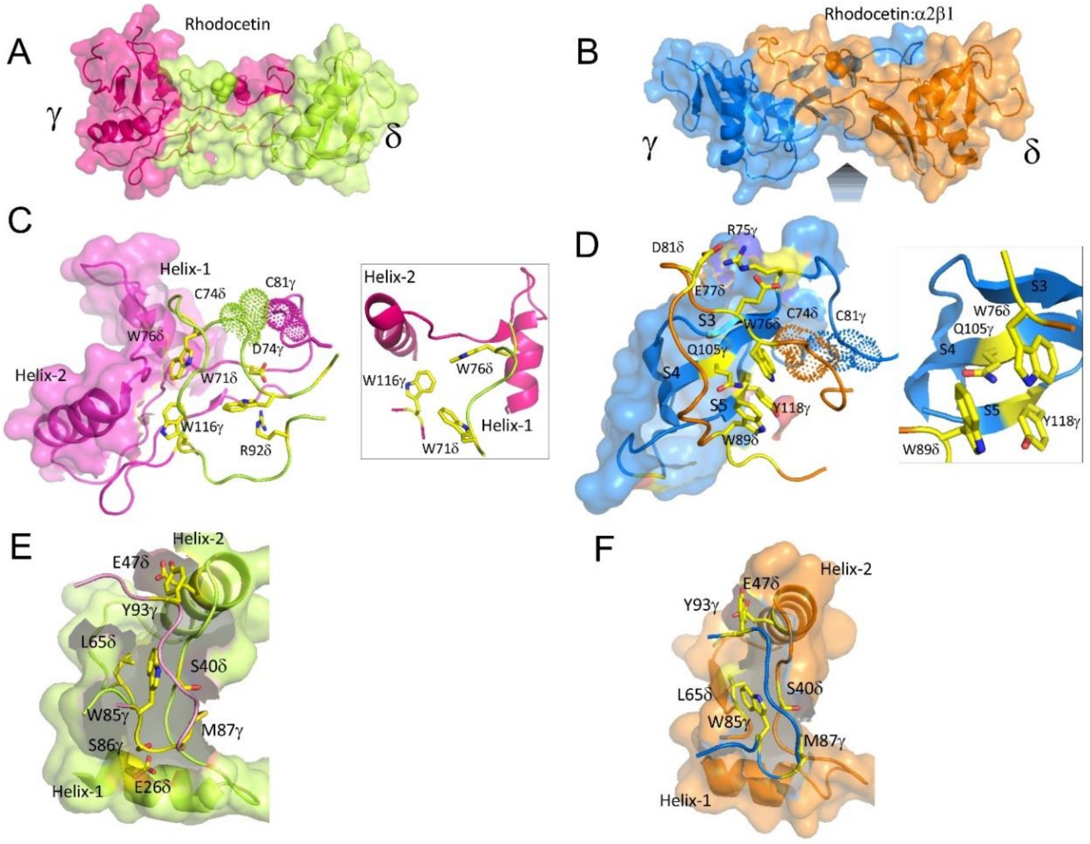


Figure 4-5. Conformational changes of RCγδ upon α2A binding.

Molecular surface presentation showing the dramatic global conformational changes which occur within the γδ subunit between the RCαβγδ tetramer (**left panels: a, c, e**) and the RCγδ-α2A complex (**right panels: b, d, f**). The RC subunits γ and δ and their conformations are color-coded red (γ) and green (δ) for RCαβγδ and blue (γ) and orange (δ) for the RCγδ-α2A complex, respectively. (**a**) and (**b**) Whereas the prospective binding face towards α2A (gray pentagon approaching from the bottom in **b**) was rather flat in RCαβγδ (**a**), RC adopted a concave surface towards α2A upon formation of the RCγδ-α2A complex (**b**). In the RCαβγδ tetramer (**c**) the loop δ - core γ interface is stabilized by the three tryptophan residues W116γ, W71δ and W76δ which form a stabilizing butterfly structure together with a salt bridge between R92δ and D74γ and a disulfide link between C81γ and C74δ which is depicted as a dotted surface. (**d**) In RCγδ-α2A, the index finger loop from subunit δ moves towards the antiparallel sheet S3-S4-S5 of subunit γ. The disulfide bridge between C81γ and C74δ is unaffected, but the stabilizing butterfly is destroyed and replaced by a hydrophobic cluster of W76δ, W89δ and Y118γ. In addition, a new salt bridge between R75γ and E77δ and D81δ is formed instead of the broken salt bridge between R92δ and D74γ. (**e, f**) A detailed view of the loop γ-core δ interaction as observed in

RCαβγδ (**e**) and RCγδ-α2A complex (**f**), respectively. In both cases the interface is highly conserved and does not alter its conformation upon the transition from RCαβγδ to RCγδ-α2A complex. The index finger loop residues of RCγ remain oriented towards the same residues of the bridge element between both helix 1 and helix 2 of the RCδ core.

The rigid body arrangement can best be described as a flipping of helices 1 and 2 between the RCγ and RCδ subunits whilst maintaining the same relative orientation of the two helices within their respective core domains. An additional consequence of this rigid body movement is a conformational shift of the connecting finger loop to track the motion of the opposing core domain. As a result, the two core domains flipped over with respect to each other and bent towards the α2A domain to form a concave binding surface such that the RCγδ residues involved in α2A binding were brought into the correct orientation for binding the α2A domain.

The apical ends of the index finger loops were in close contact with the CLRP core element of the opposite subunit, forming the two interfaces: loop γ-core δ and loop δ-core γ. Whereas the former hardly changed (**Figure 4-5e, f**), the latter showed a dramatic shift within the RCγδ-α2A complex as compared to the RCαβγδ tetramer (**Figure 4-5c, d**). In the loop δ-core γ interface of RCαβγδ tetramer (**Figure 4-5c**) a tryptophan core composed of three residues W76δ, W71δ, and W116γ together with a salt bridge between R92δ and D74γ stabilized the index finger loop of the RCδ subunit and oriented it towards the RCγ subunit core sequence connecting helices 1 and 2. However, in the RCγδ-α2A complex, the salt bridge between R92δ and D74γ found in the RCαβγδ tetramer (**Figure 4-5c**) was broken. R92δ now formed a hydrogen bond to the main chain of D219 in the α2A loop 2 and a new salt bridge was observed between R75γ and E77δ and D81δ (**Figure 4-5d**). In addition, the RCδ subunit index finger loop became embedded within the antiparallel sheet S3-S4-S5 of the RCγ core such that the indole moiety of W76δ now made van der Waals contacts to Q105γ and Y118γ (see inset **Figure 4-5c, d**).

As a result of these enormous conformational changes, especially at the loop δ core γ interface, the rigid cores of the two RCγδ subunits swung towards each other by about 40-50° around a hinge located in the center of the index finger swap domain between the cores. This global movement had two major consequences. First, as the RCδ subunit snapped into its new position the three key residues of RCγ (L66, R109, W110) underwent a local conformational change that transformed them into an orientation that is competent for α2A binding (**Figure 4-6a**). Second, as a consequence of the index finger loop tracking the movement of the RCγ subunit, the contact site between the RCα and RCγ subunits changed its three-dimensional structure due to the formation of the new salt bridge between R75γ and E77δ and D81δ (**Figure 4-5d**). Consequently, the previous interface between the RCγ subunit ($K^{77}EQQC^{81}$) and the RCα subunit ($N^{74}KQQR^{78}$) became sterically blocked[26]. The movement of the RCγ subunit would also produce steric clashes with the RCβ subunit, and it is likely the combination of these two events which resulted in the dissociation of the RCαβ subunit from its RCγδ counterpart. In contrast, the contact site within the RCδ subunit would allow integrin binding irrespective of the conformational change of RC, as their local positions and orientations remained almost unchanged (**Figure 4-6b**). In fact, the distance between Y60δ and Y94δ within the RCδ contact sites only changed slightly from 21.7 Å to 20.4 Å (Fig. 6B), while their distances towards W110γ of the RCγ contact site were reduced from 47.5 Å to 31 Å and from 28.4 Å to 18.6 Å, respectively when comparing the structure of RCαβγδ and RCγδ-α2A complex. This illustrates how significant a reorganization of the RCγδ is required to facilitate the formation of the ultimate inhibitory RCγδ-α2A complex.

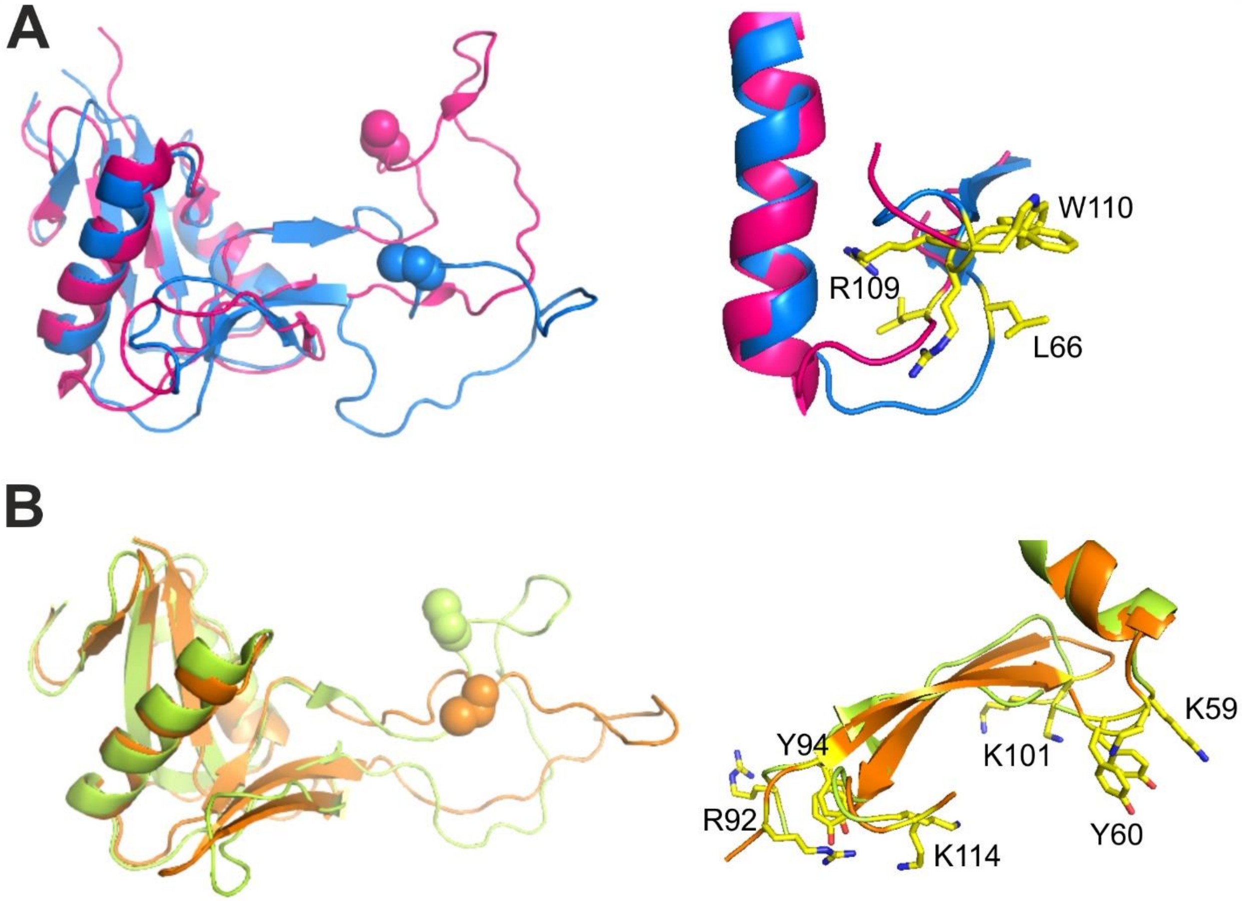


Figure 4-6. An overview of the RCγ and RCδ binding residues, depicting the local conformational changes which occur upon α2A binding.

(**a**) A comparison of the RCγ subunit binding site (L66/R109/W110) between the RCαβγδ (purple) and RCγδ-α2A complex (blue) structures. Due to the global movements within the index finger swapping domain that accompany the formation of the RCγδ-α2A complex, a local repositioning of the key α2A interacting residues within RCγ takes place such that they adopt an orientation that is compatible for α2A binding. (**b**) A comparison between the two RCδ subunit binding sites (K59/Y60/K101 and R92/Y94/K114) between the RCαβγδ (yellow) and RCγδ-α2A complex (orange) structures. In contrast to RCγ, all the RCδ residues involved in α2A binding would be in an α2A competent orientation in both the RCαβγδ (yellow) and RCγδ-α2A complex (orange) structures, with the exception of R92 which forms an internal salt bridge with D74γ in the RCαβγδ tetramer but interacts with D219 of α2A in the RCγδ-α2A complex.

Interaction of the RCγ subunit with loop2 of α2A-domain is essential for RC binding to the integrin

Unlike helix C, the docking site S^{214}QYGGD^{219} did not change its conformation between the "open" and "closed" conformation of the α2A domain. To analyze its role, we challenged RC binding to α2A with the monoclonal antibody JA202. Its epitope had previously been mapped to the sequence QTS^{214}QY[42] and thus overlapped with the RCγ subunit docking site. Among different antibodies against distinct epitopes within α2A, JA202 was the only monoclonal antibody which sterically inhibited RC binding to the α2A domain in a dose dependent manner (**Figure 4-7a**).

A comparison of integrin α2 chains from different species showed a high interspecies homology of the loop 2 sequence, S^{214}QYGGD^{219}LT^{221} (**Appendix 2, Figure II-4**). In contrast, this sequence was absent in A-domains of other integrin α subunits, suggesting that it served as a selective docking site for RC on α2β1 integrin (**Appendix 2, Figure II-5**). Therefore, we replaced the α2A sequence S^{214}QYGGD^{219}L for the corresponding sequence VGRGGRQ of the α1A-domain, and tested binding of RC to this α2A-L$2^{\alpha1}$ mutant. Although this α2A mutant was still able to bind RC, the binding affinity was reduced as indicated by an increase of the K_d-value from 0.76 ± 0.12 nM to 2.70 ± 0.39 nM (**Figure 4-7b**). In parallel to the α2A-L$2^{\alpha1}$ mutant, we exchanged residues in the loop 2 that interacted with RC (**Figure 4-7c**), specifically S214, Y216, D219 as well as the G217 and G218 that are conserved in both integrin α1 and α2 loop 2 sequences, to see which residues were functionally important for the RCγδ-α2A binding. The S214G and D219A mutants which are located at the outer edges of loop 2 gave K_d values of 0.77 ± 0.32 nM and 5.2 ± 1.36 nM, respectively, while the Y216G mutant in the centre of the loop gave a K_d value of 1.98 ± 0.64 nM (**Figure 4-7d, e**). In contrast, mutating either of the conserved glycine residues of loop 2 by generating G217K and G218L resulted in a complete loss of RC binding (**Figure 4-7d**). This result is in agreement with our structure findings (**Figure 4-7c**), which showed that anything larger than a glycine at either position 217 or 218 would sterically clash with the indole side chain of W110γ. In addition, we chemically modified the solvent exposed W110γ of RC with 2-nitro-phenylensulfenylchloride (NPS-Cl), which introduced a bulky 2-nitro-phenylensulfenyl group onto the indole side chain. The modified W110γ is no longer able to stack above the two glycines G217 and G218 causing a loss of RC binding to the α2A domain (**Figure 4-7f**). Taken together, these results demonstrated that the interaction of W110 of RCγ and the loop 2 of α2A is highly specific and essential for the formation of the high affinity and inhibitory RCγδ-α2A complex.

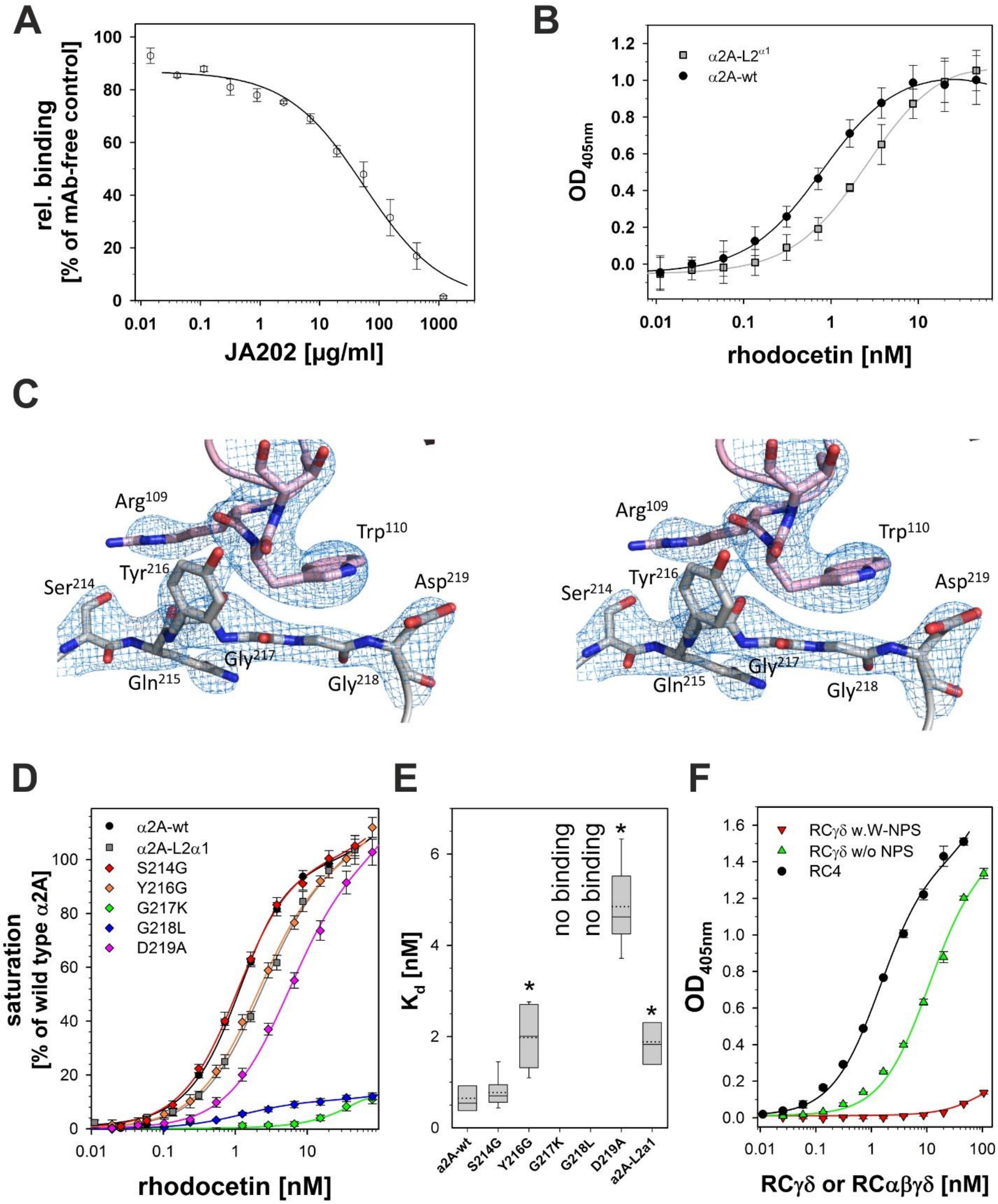

A
rel. binding
[% of mAb-free control]
100
80
60
40
20
0
0.01
0.1
1
10
100
1000
JA202 [µg/ml]
B
OD405nm
1.2
1.0
0.8
0.6
0.4
0.2
0.0
α2A-L2α1
α2A-wt
0.01
0.1
1
10
rhodocetin [nM]
C
Arg109
Trp110
Tyr216
Ser214
Asp219
Gly217
Gln215
Gly218
D
saturation
[% of wild type α2A]
α2A-wt
α2A-L2α1
S214G
Y216G
G217K
G218L
D219A
100
80
60
40
20
0
0.01
0.1
1
10
rhodocetin [nM]
E
Kd [nM]
6
4
2
0
no binding
no binding
*
*
*
a2A-wt
S214G
Y216G
G217K
G218L
D219A
a2A-L2a1
F
OD405nm
RCγδ w.W-NPS
RCγδ w/o NPS
RC4
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.01
0.1
1
10
100
RCγδ or RCαβγδ [nM]

Figure 4-7. Loop 2 of the α2A domain is the interaction site for the RCγ subunit.

(**a**) Loop 2 of α2A is an additional binding site for RC. It contains the epitope for the mAb JA202 which inhibits binding of RC to immobilized α2A. Bound RC was quantified by ELISA and values were normalized to non-inhibited controls. One set of inhibition curves out of three independent experiments with each measurement made in triplicate and the means ± s.d. for each data point shown. (**b**) The α2A loop 2 sequence was replaced with the homologous sequence VGRGGRQ of integrin α1 (α2A $L2^{\alpha1}$ mutant). The binding irrelevant antibody JA218 was immobilized to capture α2A wild-type (wt) and α2A-$L2^{\alpha1}$. They were titrated with RC and bound RC was quantified as in (**a**). One set of titration curves out of 4 independent experiments, each done in triplicates, is shown with the means ± s.d. indicated. The α2A-$L2^{\alpha1}$ mutant (light gray ■) significantly reduced affinity for RC compared to the wild type (●) (p=0.0013, two-tailed t-test) (**c**) Stereo view of the α2A loop 2 sequence in contact with the RCγ contact site. The Sigma-A weighted 2Fo-Fc map is shown at 1.5σ contour level. The two glycine residues, G217 and G218, form the bottom of a shallow dimple, which is flanked on either side by the side chains of Y216 and D219, in addition with residue N154 of loop 1 (not shown). The indole side chain of W110γ stacks directly above this dimple and interacts with the main chain of the two glycine residues. (**d**) Point mutation analysis of the α2A loop 2 sequence S^{214}QYGGD^{219}. The binding activity of these mutants for RC was tested as in (**b**). Binding signals taken from at least seven independent titration curves for each mutant were normalized to the saturation signal of wild type α2A. Means ± s.e.m. are shown for the mutants (◆ of different colors) in comparison to wild type (●) and the α2A-$L2^{\alpha1}$ mutant (light gray ■). This analysis showed that the two glycines at position 217 and 218 were key to the RCγδ-α2A interaction as only mutations abrogated α2A binding. (**e**) The K_d values of the loop 2 point mutations for binding to RC as derived from (**d**). At least seven titration curves were evaluated for each mutant. The K_d values were pairwise compared to the K_d value of the wild type α2A domain in a two-tailed Student´s t-test. Significant difference (p<0.02) is asterisked (*). (**f**) Modification of tryptophan residues of RCγδ with 2-nitrophenylensulfenylchlorid (NPS-Cl) showed that W110γ is required for α2A domain binding. The wells of a microtiter plate were coated with 10 µg/ml α2A domain and titrated with RCαβγδ (●), with non-modified RCγδ (green ▲) and with RCγδ with chemically modified W110γ (W-NPS, red ▼) One representative out three independent titration experiments done in duplicates is shown with the means ± s.d. indicated.

Discussion

Our study reveals not only the interaction sites within RC and its molecular target, the integrin α2A domain, but also the conformational changes that take place within the RCγδ subunit upon α2A binding and the relevance of the two contact sites within α2A for RCγδ binding. Moreover, these data suggest a molecular mechanism for the avid and selective interaction of this CLRP and its target.

CLRP dimers recognize other target molecules, such as factor IX/X, and the A domain of vWF by forming a bay region with their joint index finger loop swap domain and two flanking core domains. This concave face shapes the binding sites for clotting factors IX and X[43,44] and the vWF-factor A-domain[45]. Due to their importance in hemostasis, clotting factors and vWF are valid targets for CLRPs from snake venoms. Bitiscetin and botrocetin interact with the vWF-A1 domain without or together with the GPIb receptor[27,45,46]. These studies showed that these snake venom toxins can approach the A-domain from different orientations[35,45,46]. In yet another orientation, EMS16 approached the α2A domain of α2β1 integrin, which is homologous to vWF-A1 domain, along its top face directly above the metal binding site and collagen binding crevice, thus preventing collagen from binding[27]. EMS16 and RC are the two α2β1 integrin-binding CLRPs whose crystal structures in both the unliganded and the complexed with their target A-domains have been resolved so far[26,47]. Although RC approached the α2A domain in a similar orientation to EMS16, our data reveal that RC, in contrast to any known CLRP structure[27,45,46], undergoes a dramatic conformational change to form a concave binding surface. The heterodimeric EMS16 did not alter its molecular structure upon α2A binding[27,47], as the concave binding surface required for α2A binding was already preformed. This difference in mode of α2A binding between EMS16 and RC is either determined by the distinct quaternary structures of the dimeric EMS16 versus the tetrameric RC and/or by the different purification protocols. When we employed the same purification procedure for RC as for EMS16 and other CLRPs[28-30,48] using reversed phase chromatography performed in 0.1 % TFA solution, the RC tetramer dissociated into its subunits α, β and γδ[49]. The RCγδ subunit alone was still able to bind α2A and to block α2β1 integrin-mediated platelet aggregation specifically[50]. Only when applying a milder purification protocol could we obtain a stable RC tetramer and the RCγδ-α2A complex whose different conformational structures are presented here.

Our crystal structure of the RCγδ-α2A complex reveals a geometry of interaction similar to the α2A bound EMS16, suggesting that the α2β1 integrin blocking CLRPs may have a more uniform binding mechanism than the vWF binding CLRPs (**Figure 4-8**). Both CLRPs share the same two contact sites within the α2A domain: the conformationally stable loop 2 sequence (**Figure 4-8c**) and the helix C of the “closed” conformation (**Figure 4-8d**). Helix C is recognized by the structurally robust contact area of the RCδ subunit or the homologous EMS16 subunit β (or B). Apart from slight variations of the K59δ side chain and the loop 2 Y216 side chain (**Figure 4-8d**) adopting an alternate conformation to form a hydrophobic interaction with L66γ, the structures of both complexes are almost identical in this region. In our studies, the role of the loop 2 sequence $S^{214}QYGGD^{219}$ was reinforced by the JA202 antibody, whose epitope overlaps with this loop 2 sequence and inhibited RC binding completely, presumably due to steric hindrance by the bulky antibody. More subtly, recombinant exchange of the respective loop 2 sequence with the homologous sequence of integrin α1 showed that the loop 2 sequence changes the affinity of the venom component towards the integrin α2 subunit. Similar reductions in the affinity of RC for α2A were also observed with the loop 2 mutants Y216G and D219A.

However, a loss of binding was obtained with the G217K and G218L mutants. These two glycine residues form part of a shallow dimple on the α2A surface which is covered by W110 of RCγ subunit. In the molecular structure of the RCγδ-α2A complex, there is not any space to accommodate anything larger than a glycine at either of these two positions which explains the loss of function of these two mutants. The loop 2 sequence of the integrin α2A domain is evolutionary conserved between different animal species, especially the GG motif at positions 217 and 218, but varies remarkably between other integrin α subunits. This suggests that rhodocetin´s specificity is mediated by the integrin α2 specific loop 2 sequence, as RC affects α2β1 integrin mediated platelet blockage in various potential preys, but does not affect biological functions mediated by other integrins. Our conclusion, that this cluster of RCγ W110 and G217/G218 of the α2A loop 2 sequence is a key to the RCγδ-α2A interaction is further supported by the fact that the RC binding is completely lost if the bulky chemical adduct of 2-nitrophenylsulfenyl is introduced to the indole side chain. It is noteworthy, that the loop 2 sequence is also relevant for collagen binding as it forms a hydrophobic contact for the phenylalanine side chain of the middle strand of the trimeric integrin recognition motif of collagen[15], albeit not as close a contact as with the RCγ W110 side chain.

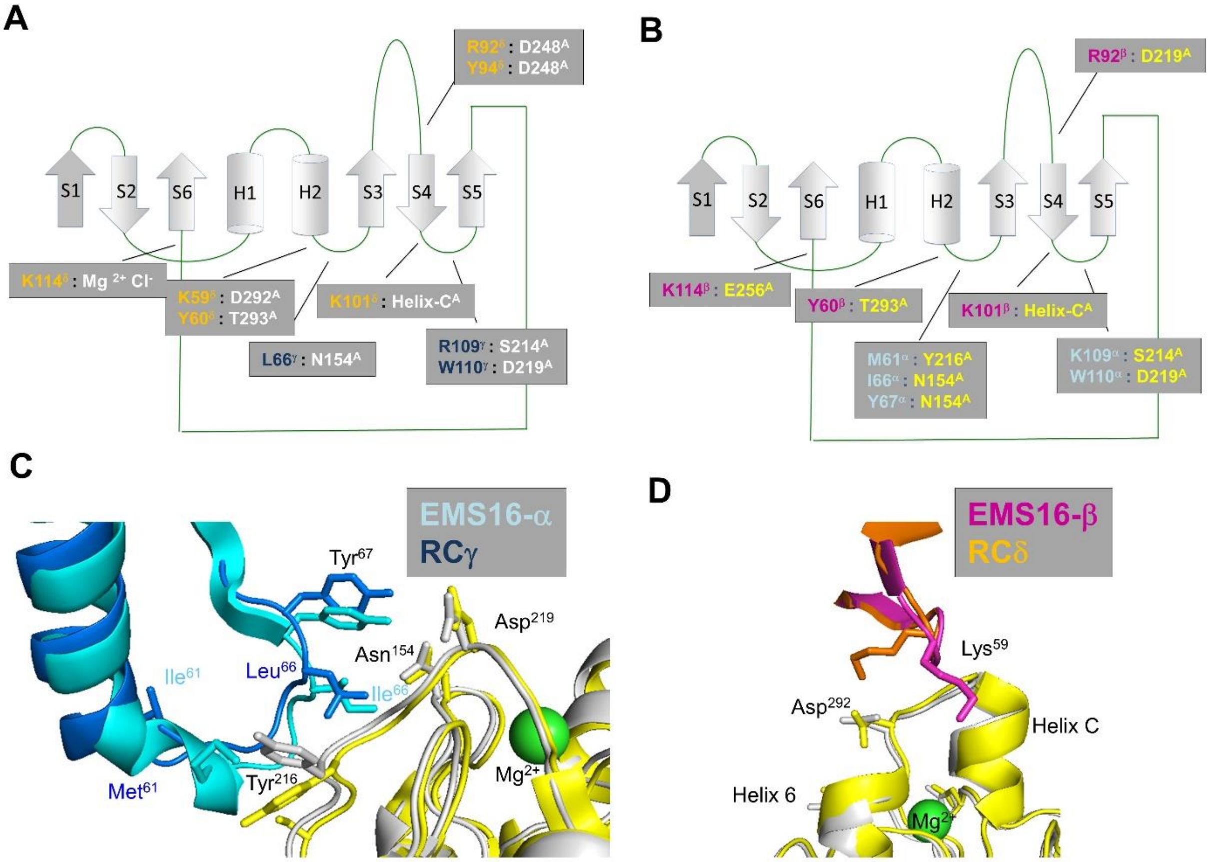


Figure 4-8. A comparison of the RCγδ-α2A and EMS16αβ-α2A binding interfaces.

(**a, b**) The CLRP folds of both homologous subunits of RCγδ (**a**) and EMS16αβ (**b**) are highly homologous with many of the residues involved in the α2A binding conserved between the two proteins. These residues have been mapped onto the CLRP fold and color coded for RC (blue and orange for the γ and δ subunits, respectively, in (**a**)) and for EMS16 (light blue and magenta for the α and β subunits, respectively, in (**b**)). The partnering residues of the α2A domain contacted by RC and EMS16 are color-coded in white and yellow, respectively. The same color coding scheme is used throughout the figure. (**c, d**) A superposition of the key residues from RCγ/EMS16α at the loop 2 binding site (**c**) and of RCδ/EMS16β at the helix C binding site (**d**), respectively, on α2A. The contact sites are largely conserved between RCγδ/EMS16αβ and α2A although there are a couple of notable differences. For example, L66 of RCγ contacts Y216 of α2A in addition to the N154 of loop 1 observed for the corresponding I66 of EMS16α. In addition, K59 of RCδ forms a salt bridge to D292 of α2A whereas in EMSβ the corresponding K59 points towards helix C.

Based on our findings we suggest the following mode of action (**Figure 4-9**). RCαβγδ interacts with helix C of the α2A domain through the RCδ subunit, where the interacting residues are already in binding competent orientation. This stabilizes the “closed” conformation of α2A. As a consequence of the movement of RCγ, the RCαβγδ tetramer changes conformation such that RCαβ dissociates from the heterotetrameric assembly. Coupled to this dissociation is the reorganization of Leu66, Arg109 and Trp110 of RCγ to interact with loop 2 sequence $S^{214}QYGGD^{219}$. Having established both interaction sites, RCγδ firmly binds to α2A and holds it in the “closed” conformation, thereby blocking collagen binding and antagonistically turning off α2β1 integrin signaling. After its release upon formation of the high affinity RCγδ-α2β1 complex, the RCαβ subunit plays another important role in blocking GPIb and consequently vWF induced platelet aggregation[49]. Moreover, our biochemical data showed that the RCαβ subunit is significantly more soluble than the RCγδ subunit[40]. Therefore, it likely acts as a solubility enhancer to ensure that the RCγδ subunit is delivered to α2β1 integrin. Once RCγδ has bound to its target and the RCαβ subunit has been released, RC effectively shuts down the two platelet receptors, α2β1 integrin and GPIb, thereby effectively blocking both collagen induced and vWF induced platelet activation and aggregation.

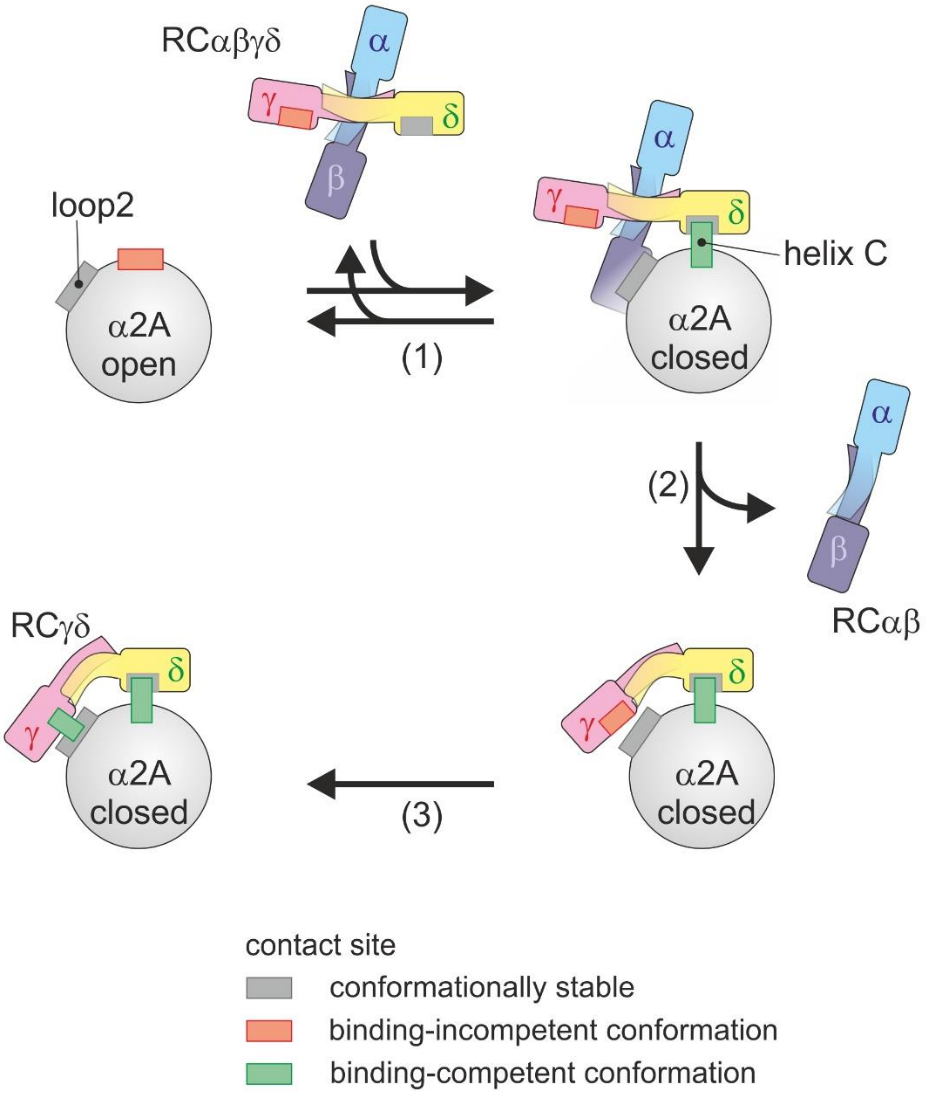


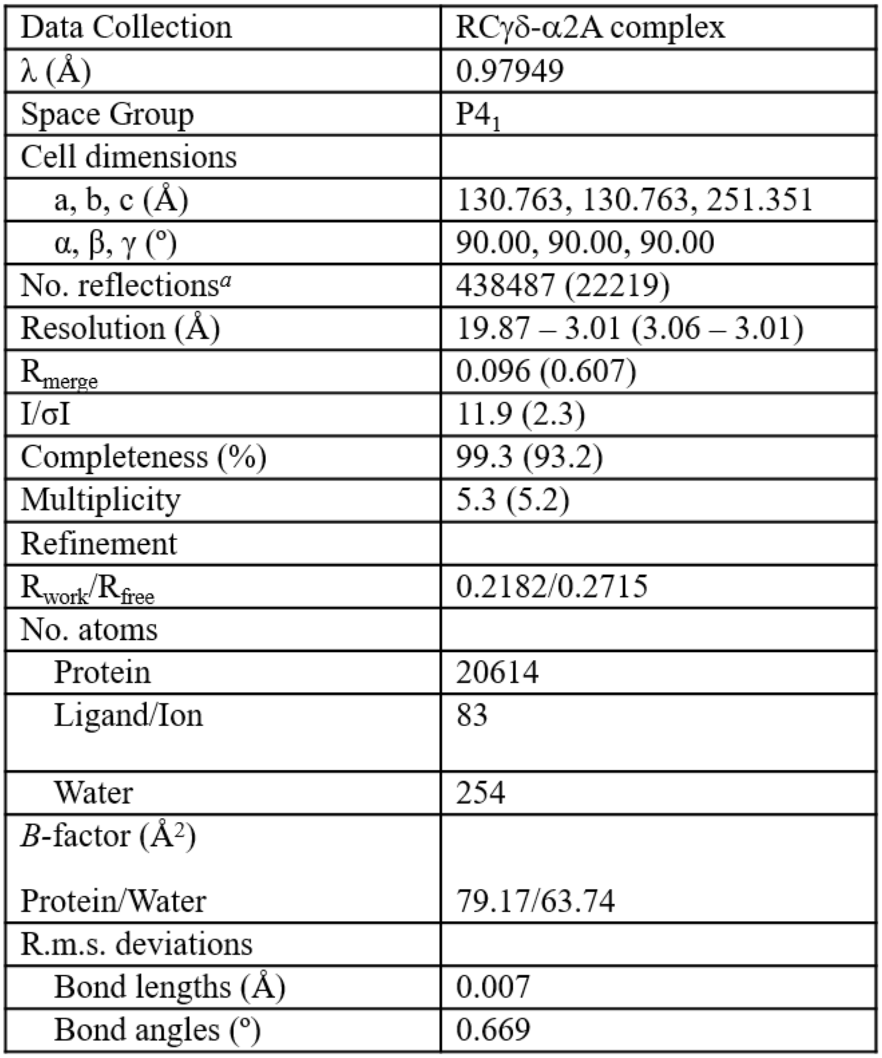
Figure 4-9. Molecular mechanism of the RCγδ-α2A interaction.

As RCαβγδ binds to α2A in its "closed" conformation, it induces the conformational change of α2A from its "open" to "closed" conformation and thus shifts the conformational equilibrium (1). This interaction is mediated via the conformationally robust RCδ interaction site within helix C, which is only present in the "closed" conformation of α2A. Subsequently, the index finger loop of RCγ changes its conformation which is accompanied by a global movement of both RC core domains towards each other and by a release of the RCαβ subunit (2). As the RCαβ subunit diffuses away, this step is likely irreversible in nature. The global shape change of RCγδ forms a new bay region which embraces α2A and locally leads to the repositioning of RCγ key residues, which forms another binding competent interacting site in RCγ for the α2A loop 2 (3).

In summary, a comparison of the RCγδ-α2A structure with the EMS16-α2A integrin complex[27] shows that the residues involved in the binding of RC and EMS16 to α2β1 integrin are highly conserved, which suggests that the binding site is conserved at least for these two snake venom CLRPs even if the way by which it is formed is dramatically different. Unique to the formation of the RCγδ-α2A complex are the substantial conformational changes in both binding partners, whereby the two interaction sites between RC and α2A have a common characteristic feature. At both interaction sites, a conformation-dependent binding site (α-helix C of α2A, and the RCγ chain contact site) interacts with a structurally robust counter-site (RCδ chain contact site, and loop 2 sequence $S^{214}QYGGD^{219}$, respectively). The conformational changes are mutually induced upon RCγδ-α2β1 complex formation suggesting a cooperative effect of CLRP and target receptor, which is responsible for the high affinity and specificity of the interaction. The proposed molecular mechanism and the key contacts reported here may be instrumental for the design of novel α2β1 integrin inhibitors.

Tables

Table 4-1. Data and refinement statistics of the RCγδ-α2A crystal structure



Data Collection	RCγδ-α2A complex
λ (Å)	0.97949
Space Group	$P4_1$
Cell dimensions	
a, b, c (Å)	130.763, 130.763, 251.351
α, β, γ (°)	90.00, 90.00, 90.00
No. reflections[a]	438487 (22219)
Resolution (Å)	19.87 – 3.01 (3.06 – 3.01)
R_{merge}	0.096 (0.607)
I/σI	11.9 (2.3)
Completeness (%)	99.3 (93.2)
Multiplicity	5.3 (5.2)
Refinement	
R_{work}/R_{free}	0.2182/0.2715
No. atoms	
Protein	20614
Ligand/Ion	83
Water	254
B-factor (Å^2) Protein/Water	79.17/63.74
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.669

[a] Statistics of the highest resolution shell are shown in parenthesis

Materials and Methods

Materials

Rhodocetin (RC) and its γδ-subunit were isolated as previously described[40,51].The monoclonal antibodies (mAbs) against rhodocetin, among them IIIG5 from mouse and IC3 from rat, were generated and isolated as previously described[40]. The murine mAbs against the human α2A domain, JA202 and JA218 were a generous gift from D. Tuckwell (formerly University of Manchester, UK)[40,42]. PCR-primers were obtained from Eurofins (Eurofins Genomics, Germany) and are written in 5´-3´ direction. Restriction enzymes and molecular biology reagents were from Thermo Fisher Scientific (Germany) unless otherwise stated. Cloning products and expression vectors were validated by DNA sequencing (Eurofins Genomics).

Tryptophan-specific chemical modification of RC

RC, dissolved at 110 µM in 30% acetic acid solution was treated with 9.2 mM 2-nitrophenylensulfenylchloride (NPS-Cl, TCI Chemicals, Germany) or left untreated for 1 h at 20°C in the dark according to[52], subsequently dialyzed against 0.1 % trifluoroacetic acid (RP-solution) and separated on a Supercosil C18 column (Supelco, Germany) by reversed phase chromatography as described[26]. The RCγδ containing fractions were pooled, lyophilized and dissolved in RP-solution, containing 30 % acetonitrile. Purity was assessed by SDS-PAGE. Spectroscopic evaluation at 365 nm according to[52] confirmed the covalent modification of RC tryptophan residues with 2-nitro-phenylensulfenyl (NPS)-groups.

Isolation of RCγδ-α2A complex

The His_6-tagged α2A domain was generated as previously described[26,53]. It was loaded onto a HiTrap Ni Sepharose column (GE Healthcare; 5 ml) previously equilibrated with $PBS/MgCl_2$-buffer, pH 7.4 (20 mM sodium phosphate, pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$). After washing with the same buffer, the RCαβγδ containing fractions from the rhodocetin isolation with MonoS column[51] were applied to the α2A domain loaded Ni Sepharose column, after having been treated with 0.5 µM PMSF and 1 µg/ml aprotinin to prevent proteolytic digestion by potentially contaminating snake proteases. After RCαβγδ had bound to the Ni Sepharose immobilized α2A domain, the HiTrap Ni Sepharose column was washed with $PBS/MgCl_2$-buffer, pH 7.4. Then, the column was washed with PBS/EGTA-buffer, pH 7.4 (5 mM EGTA in 20 mM sodium phosphate, pH 7.4, 150 mM NaCl) and the RCαβ subunit eluted. After another washing step with $PBS/MgCl_2$-buffer, pH 7.4, the RCγδ-α2A-complex was eluted with a linear gradient of 0-200 mM imidazole in $PBS/MgCl_2$-buffer, pH 7.4 buffer from the HiTrap Ni Sepharose column. Protein concentration in the imidazole eluate was determined using the Bradford reagent (BioRad). For crystallization, the complex containing fractions were pooled and digested with TPCK-treated trypsin (Sigma-Aldrich) at an enzyme-substrate ratio of 1:100 at 37°C for 1 h. The digest was stopped with 1 mM PMSF, concentrated and separated by gel filtration to remove excess α2A domain, trypsin and contaminating peptides from the RCγδ-α2A complex. The TSK G2000SWXL chromatography was performed in 10 mM HEPES, pH 7.4, 100 mM NaCl buffer. The RCγδ-α2A complex was concentrated by ultrafiltration and its protein concentration determined with the Bicinchoninic Acid Protein Assay (BCA, Thermo Fisher Scientific). To analytically prove, the physical contact of both partners, the complex was cross-linked with 0.5 mM bi-sulfosuccinimidyl-suberate (BS^3, Thermo Fisher Scientific). Its IEP was determined to be pH 6.5-6.8 and pH 6.7 by isoelectric focussing in precast ZOOM pH 3-10 gels (Thermo Fisher Scientific) and by analytical chromatofocussing on a MonoP column (GE HealthCare) in a pH gradient of 7.4 to 4.0, respectively.

Crystallization, data processing and structure refinement

Crystals of 10 mg of RCγδ-α2A were grown by hanging-drop vapor diffusion at 293 K by mixing 2 µL of protein solution with 2 µL reservoir solution containing 2.65M ammonium sulfate and 100 mM Tris pH 8.0. Crystals appeared after 6 weeks and were soaked in mother liquor containing 20% glycerol for 5-10 min before flash freezing in liquid nitrogen. Diffraction data was collected at the Canadian Light Source CMCF-08ID-1 beamline (λ=0.97949Å) at 100 K using a Rayonix MX225 CCD detector. The dataset was indexed, integrated and scaled with MOSFLM[54] and the CCP4-package[55]. The spacegroup is $P4_1$ with 6 molecules in the asymmetric unit (see **Table 1**). The phases were determined by rigid body refinement using the previously solved rhodocetin structure (PDB code 3GPR) in Refmac[56,57]. The model was built and refined without NCS restraints using Coot[58] and refined with the Phenix software package[59]. The crystallographic data and refinement statistics are summarized in Table 1. The final coordinates and structure factor amplitudes were deposited in the PDB (PDB: 5THP).

Generation of integrin α2 A-domain mutants

The human α2A domain and its mutants were produced in a bacterial expression system. The expression vectors encoding the disulfide-locked conformation mutants of α2A were generated using a previously described pET15b-His$_6$-α2A construct (residues 142 through 337 of human integrin α2). To replace the endogenous cysteine residues at 150 and 270, this plasmid was used as template for a two step-PCR with the three primer pair sets (i) HT*fwd*(CTCTCCATGGGCTCTTCTCATCATCATCATCATCATTC) and R1(C11A) (CATC**AGC**CACAACCACAAC), (ii) F2(C11A) (TTGTG**GCT**GATGAATCAAATAG) and R2(C131A) (TTG**GCT**TGATCAATCACAGC), and (iii) F3(C131A) (ATTGATCAA**GCC**AACCATGAC) and α2A*rev* (CGGACATATGCTAACCTTCAATGCTGAAAAATTTG), in the first set of reactions. The three amplicons were purified and again PCR-amplified with the outer primer pair HT*fwd* and α2A*rev* to a 670 bp amplicon, which after A-tailing with Taq DNA polymerase was intermediately ligated into pCR®2.1 TOPO-TA, excised with *Nde*I and *Nco*I and the restriction fragment was subcloned into the linearized, *Nde*I, *Nco*I-cleaved pET-15b expression vector. The final expression plasmid pET-15b-His$_6$-α2A(C150,270A), was transformed into *E. coli* BL21 (DE3).

To generate the disulfide-locked conformation mutants of α2A, which share the same K168C mutation but differ in E318C (open conformation: K168C, E318C) or A325C (closed conformation: K168C, A325C) three rounds of PCR amplification were performed. In the first site-directed mutagenesis K168C was introduced by amplifying the entire plasmid with the back-to-back primer pair K168C fw (AAGGCCTGGATATAGGCCCC) and K168C rev (GTACAAAGCATTCCAAAAAATTCTTTACTGC). Based on this mutation the final two mutants (K168C, E318C; K168C, A325C) were similarly generated using the primer pairs E318C fw (GTCTGATTGCGCAGCTCTACTAGAAAAG) / E318C rev (ACATTGAAAAAGTATCTTTCTGTTGGAATAC) and A325C fw (ATTAGGAGAACAAATTTTCAGCATTGAAG) / A325C rev (GTCCCGCACTTTTCTAGTAGAGCTG). For each site directed mutagenesis only one primer contained the specific mutation. The PCR products were amplified by the Phusion Hot Start II polymerase and covered the whole template vector (6307 bp) with the mutation. After the original, methylated vector had been digested with *DpnI,* the amplicons were purified using the DNA Clean & Concentrator Kit (Zymo Research), followed by 5´-phosphorylation with T4 polynucleotide kinase and religated using T4 DNA ligase. For protein expression, *E. coli* strain BL21 (DE3) were transformed with the validated plasmid constructs encoding the α2A domain in its open (pET-15b-His$_6$-α2A-C150/270A-K168C/E318C) and closed (pET-15b-His$_6$-α2A-C150/270A-K168C/A325C) conformation.

The α2A-L2$^{\alpha1}$ mutant, in which the sequence S^{214}QYGGDL is replaced by the corresponding loop 2 sequence V^{214}QRGGRDQ of the integrin α1 A-domain, was generated by two step-PCR. The pET15b-construct encoding the His-tagged α2A domain[26] was used as a template. The following primer pairs: α2A fw (GGATATCTGCAGAATTCGCCCTTC) and R1_ a1insert into a2 (CTTTACTAACATCGTTGTAGGGTCTGTCACGTCGCGCCACCAGCGGTC), and F1_a1insert into a2 (GTGCAGCGCGGTGGTCGCCAGACAAACACATTCGGAGCAATTC) and α2A rev (AGGCCATATGCTAACCTTCAATGCTGAAAATTTG) amplified the N and C-terminal half of the cDNA. The two amplicons were mixed and amplified with the outer primer pair. The resulting 680 bp-amplicon was trimmed with *Nco*I and *Nde*I, ligated into a correspondingly cut pET-15b vector, verified by sequencing and transformed into *E. coli* BL21(DE3).

Point mutations within the loop 2 sequence were also generated by a two step-PCR using the wild-type α2A-encoding cDNA as template. First, cDNA-fragments encoding the N- and C-terminal halves of α2A were amplified by using the two pairs of forward outer and reverse inner primers and of forward inner and reverse outer primers, respectively, as summarized in Appendix 2, Table II-1.

The amplicons were purified and taken as template for a second PCR with the outer primer pair to obtain the wild-type and mutant α2A domains encoding cDNAs, which were digested with *NdeI* and *BamHI* and ligated into the likewise cut pET-15b vector. After verification by sequencing, the expression vectors were transformed into *E.coli* BL21(DE3). All α2A domain mutants were purified using HiTrap Ni Sepharose column (GE HealthCare) as per the wild type.

Binding and inhibition assays of α2A-domain with RC

The wells of a half area microtiter plate (Costar) were coated with 10 µg/ml His-tagged α2A domain in TBS/Mg buffer (50 mM Tris/HCl, pH 7.4, 150 mM NaCl, 3 mM $MgCl_2$) at 4°C overnight. After washing twice with TBS/Mg buffer, the wells were blocked with 1% BSA in TBS, pH 7.4, 2 mM $MgCl_2$ for 1 h at room temperature. The immobilized α2A domain was titrated with a serial dilution of RCαβγδ or RCγδ without and with NPS-modified tryptophans in the blocking buffer for 1.5 h. For the mAb inhibition experiment, RC at a constant concentration of 2 nM was added to the wells in either absence or presence of mAb JA202 against RC. After washing twice with HEPES-buffered saline (HBS) (50 mM HEPES/NaOH, pH7.4, 150 mM NaCl, 2 mM $MgCl_2$), bound RC was fixed with 2.5 % glutaraldehyde in the same solution for 10 min at room temperature. After three additional washes with TBS/Mg buffer bound RC was quantified by ELISA using a primary rabbit antiserum against RC and a secondary alkaline phosphatase conjugated anti rabbit-IgG antibody, each diluted 1:2000 in 1% BSA/TBS/Mg. Conversion of para-nitrophenyl phosphate (pNpp) to para-nitrophenolate was stopped with 1.5 M NaOH and measured at 405 nm. The titration curves were evaluated as described below. The inhibition curves were approximated by GraphPad Prism software using the inhibition vs. log [inhibitor]-approximation. To compare independent inhibition and binding experiments, the dynamic ranges were normalized to the mAb-free control and to the saturation value of the wild type α2A domain, respectively.

Alternatively, the α2A domains, either wild type or mutants, were captured using the mAb JA218 at a ligand-binding irrelevant epitope, thereby avoiding any conformational changes due to adsorption to the plastic. To this end, 2.5 µg/ml JA218 was immobilized to a microtiter well at 4°C overnight. After the wells were washed twice with TBS/Mg buffer, wells were blocked with 1% BSA in the same buffer for 1 h and then the α2A domain added at 10 µg/ml for 1 h. After washing the wells, RC was titrated and detected as described above.

Capturing ELISA with IIIG5

The mAb IIIG5 was coated to the wells of a microtiter plate at 3 µg/ml in TBS/Mg buffer overnight. After two washing steps, wells were blocked with 1 %BSA in TBS/Mg buffer for 1 h and then titrated with either RCαβγδ, RCγδ or RCγδ-α2A complex for 1.5 h at room temperature. Bound RC was fixed and quantified as described above. A mathematical approximation of the titration curve, including determination of K_d-values, is described below.

Isolation of IIIG5 epitope and mass spectrometry (MS)

IIIG5 was immobilized to cyanogen bromide activated sepharose according to the manufacturer´s instruction (GE Healthcare). RCαβγδ-containing fractions from the Mono S purification of *Calloselasma rhodostoma* venom[26] were reduced with 4 mM Tris(hydroxymethyl)phosphine (THP, Calbiochem) for 20 min at 60°C, and free thiol groups were alkylated with 16 mM iodoacetic acid. The protein was precipitated with trichloroacetic acid, washed with acetone twice, resuspended in 87.5 mM sodium bicarbonate/0.5 M urea and digested with TPCK-trypsin for 23 h at 37°C. After addition of 1 mM phenylmethylsulfonyl fluoride (PMSF), the digest was diluted with TBS/HCl buffer, pH 7.4 and loaded onto the IIIG5 column. The RC-peptide harboring the IIIG5 epitope was eluted in a pH gradient from pH 7.5 to 3.0 in 20 mM citrate buffer and further purified by reversed phase on a Supercosil C18 column in a 0-28 % acetonitrile gradient in 0.1%TFA/water. Lyophilized HPLC fractions were dissolved in 40 % methanol containing 0.5 % formic acid and analyzed by nano electrospray ionization (nanoESI) MS and MS/MS. Peptide structures were deduced from the corresponding fragment ion spectra. NanoESI MS experiments were carried out by using a SYNAPT G2-S mass spectrometer (Waters, Manchester, UK) equipped with a Z-spray source in the positive ion sensitivity mode. Typical source parameters were: source temperature: 80°C, capillary voltage: 0.8 kV, sampling cone voltage: 20 V, and source offset voltage: 50 V. For low energy collision induced dissociation (CID) experiments, the peptide precursor ions were selected in the quadrupole analyser, subjected to ion mobility separation (IMS; wave velocity 850 m/s, wave height 40 V, nitrogen gas flow rate 90 ml/min, and helium gas flow rate 180 ml/min), and fragmented in the transfer cell using a collision gas (Ar) flow rate of 2.0 ml/min and collision energies up to 100 eV (E_{lab}).

Mathematical evaluation of titration curves

In titration curves, a signal S, usually the extinction at 405 nm caused by the alkaline phosphatase-catalyzed conversion of pNpp, is measured in response to the total concentration c_0 of added titrant. Based on a Michaelis-Menten-like binding mechanism, we deduced the following equation to approximate titration curves, if the signal S and the total concentration c_0 of added ligand (RC) is known:

Equation 4-1

$$S(c_0) = (S_M - S_m) \cdot \left(\frac{(c_0 + \ c_R + K) - \sqrt{(c_0 + \ c_R + K)^2 - 4 \cdot c_0 \cdot c_R}}{2 \cdot c_R} \right) + S_m + B \cdot c_0$$

with S_M and S_m, maximum and minimum signals, respectively; c_R, the concentration of ligand binding site (equals the receptor concentration for monovalent receptors), and K, the dissociations constant K_d. The term $B \cdot c_0$ takes into account a linear change in the signal due to non-specific binding of the ligand. The five parameters S_M, S_m, c_R, K, and B are calculated by non-linear regression from titration curves.

Statistical analysis.

The data from titration and inhibition curves were statistically evaluated using GraphPad Prism software. Values were usually compared with the values obtained for the wild type α2A or non-modified RC with Student´s t-test, where the significance level was set at 1%, unless otherwise stated.

References

1. Iwamoto DV, Calderwood DA. Regulation of integrin-mediated adhesions. Curr Opin Cell Biol. 2015;36:41-7. doi: 10.1016/j.ceb.2015.06.009. PubMed PMID: 26189062; PubMed Central PMCID: PMCPMC4639423.

2. Maartens AP, Brown NH. Anchors and signals: the diverse roles of integrins in development. Curr Top Dev Biol. 2015;112:233-72. doi: 10.1016/bs.ctdb.2014.11.020. PubMed PMID: 25733142.

3. Madamanchi A, Santoro SA, Zutter MM. alpha2beta1 Integrin. Adv Exp Med Biol. 2014;819:41-60. doi: 10.1007/978-94-017-9153-3_3. PubMed PMID: 25023166.

4. Nuyttens BP, Thijs T, Deckmyn H, Broos K. Platelet adhesion to collagen. Thromb Res. 2011;127 Suppl 2:S26-9. doi: 10.1016/S0049-3848(10)70151-1. PubMed PMID: 21193111.

5. Hynes RO. Cell-matrix adhesion in vascular development. J Thromb Haemost. 2007;5 Suppl 1:32-40. doi: 10.1111/j.1538-7836.2007.02569.x. PubMed PMID: 17635706.

6. Yeh YC, Lin HH, Tang MJ. A tale of two collagen receptors, integrin beta1 and discoidin domain receptor 1, in epithelial cell differentiation. Am J Physiol Cell Physiol. 2012;303(12):C1207-17. doi: 10.1152/ajpcell.00253.2012. PubMed PMID: 23015544.

7. Zeltz C, Gullberg D. The integrin-collagen connection--a glue for tissue repair? J Cell Sci. 2016;129(4):653-64. doi: 10.1242/jcs.180992. PubMed PMID: 26857815.

8. Gillberg L, Berg S, de Verdier PJ, Lindbom L, Werr J, Hellstrom PM. Effective treatment of mouse experimental colitis by alpha 2 integrin antibody: comparison with alpha 4 antibody and conventional therapy. Acta Physiol (Oxf). 2013;207(2):326-36. doi: 10.1111/apha.12017. PubMed PMID: 23009282.

9. Peters MA, Wendholt D, Strietholt S, Frank S, Pundt N, Korb-Pap A, et al. The loss of alpha2beta1 integrin suppresses joint inflammation and cartilage destruction in mouse models of rheumatoid arthritis. Arthritis Rheum. 2012;64(5):1359-68. doi: 10.1002/art.33487. PubMed PMID: 22083543.

10. Naci D, Vuori K, Aoudjit F. Alpha2beta1 integrin in cancer development and chemoresistance. Semin Cancer Biol. 2015;35:145-53. doi: 10.1016/j.semcancer.2015.08.004. PubMed PMID: 26297892.

11. Marjoram RJ, Li Z, He L, Tollefsen DM, Kunicki TJ, Dickeson SK, et al. alpha2beta1 integrin, GPVI receptor, and common FcRgamma chain on mouse platelets mediate distinct responses to collagen in models of thrombosis. PLoS One. 2014;9(11):e114035. doi:

10.1371/journal.pone.0114035. PubMed PMID: 25415203; PubMed Central PMCID: PMCPMC4240667.

12. Momic T, Katzhendler J, Shai E, Noy E, Senderowitz H, Eble JA, et al. Vipegitide: a folded peptidomimetic partial antagonist of alpha2beta1 integrin with antiplatelet aggregation activity. Drug Des Devel Ther. 2015;9:291-304. doi: 10.2147/DDDT.S72844. PubMed PMID: 25609915; PubMed Central PMCID: PMC4294129.

13. Nissinen L, Pentikainen OT, Jouppila A, Kapyla J, Ojala M, Nieminen J, et al. A small-molecule inhibitor of integrin alpha2 beta1 introduces a new strategy for antithrombotic therapy. Thromb Haemost. 2010;103(2):387-97. doi: 10.1160/TH09-06-0358. PubMed PMID: 20126829.

14. Zhang L, Zhang C, Sun Y. Biomimetic design of platelet adhesion inhibitors to block integrin alpha2beta1-collagen interactions: II. Inhibitor library, screening, and experimental validation. Langmuir. 2014;30(16):4734-42. doi: 10.1021/la4046012. PubMed PMID: 24697658.

15. Emsley J, Knight CG, Farndale RW, Barnes MJ, Liddington RC. Structural Basis of Collagen Recognition by Integrin α2β1. Cell. 2000;101(1):47-56. doi: 10.1016/s0092-8674(00)80622-4.

16. Emsley J, King SL, Bergelson JM, Liddington RC. Crystal structure of the I domain from integrin alpha2beta1. J Biol Chem. 1997;272(45):28512-7. PubMed PMID: 9353312.

17. Shimaoka M, Lu C, Palframan RT, von Andrian UH, McCormack A, Takagi J, et al. Reversibly locking a protein fold in an active conformation with a disulfide bond: integrin

alphaL I domains with high affinity and antagonist activity in vivo. Proc Natl Acad Sci U S A. 2001;98(11):6009-14. doi: 10.1073/pnas.101130498. PubMed PMID: 11353828; PubMed Central PMCID: PMC33413.

18. Calvete JJ. The continuing saga of snake venom disintegrins. Toxicon. 2013;62:40-9. doi: 10.1016/j.toxicon.2012.09.005. PubMed PMID: 23010163.

19. Calvete JJ, Marcinkiewicz C, Monleon D, Esteve V, Celda B, Juarez P, et al. Snake venom disintegrins: evolution of structure and function. Toxicon. 2005;45(8):1063-74. doi: 10.1016/j.toxicon.2005.02.024. PubMed PMID: 15922775.

20. Clemetson KJ. Snaclecs (snake C-type lectins) that inhibit or activate platelets by binding to receptors. Toxicon. 2010;56(7):1236-46. doi: 10.1016/j.toxicon.2010.03.011. PubMed PMID: 20350564.

21. Marcinkiewicz C. Applications of snake venom components to modulate integrin activities in cell-matrix interactions. Int J Biochem Cell Biol. 2013;45(9):1974-86. doi: 10.1016/j.biocel.2013.06.009. PubMed PMID: 23811033; PubMed Central PMCID: PMC3774133.

22. Calderon LA, Sobrinho JC, Zaqueo KD, de Moura AA, Grabner AN, Mazzi MV, et al. Antitumoral activity of snake venom proteins: new trends in cancer therapy. Biomed Res Int. 2014;2014:203639. doi: 10.1155/2014/203639. PubMed PMID: 24683541; PubMed Central PMCID: PMCPMC3943284.

23. McCleary RJ, Kini RM. Non-enzymatic proteins from snake venoms: a gold mine of pharmacological tools and drug leads. Toxicon. 2013;62:56-74. doi: 10.1016/j.toxicon.2012.09.008. PubMed PMID: 23058997.

24. Vogtle T, Cherpokova D, Bender M, Nieswandt B. Targeting platelet receptors in thrombotic and thrombo-inflammatory disorders. Hamostaseologie. 2015;35(3):235-43. doi: 10.5482/HAMO-14-10-0049. PubMed PMID: 25634564.

25. Arlinghaus FT, Eble JA. The collagen-binding integrin alpha2beta1 is a novel interaction partner of the Trimeresurus flavoviridis venom protein flavocetin-A. J Biol Chem. 2013;288(2):947-55. doi: 10.1074/jbc.M112.399618. PubMed PMID: 23204528; PubMed Central PMCID: PMC3543044.

26. Eble JA, Niland S, Bracht T, Mormann M, Peter-Katalinic J, Pohlentz G, et al. The alpha2beta1 integrin-specific antagonist rhodocetin is a cruciform, heterotetrameric molecule. FASEB J. 2009;23(9):2917-27. doi: 10.1096/fj.08-126763. PubMed PMID: 19369383.

27. Horii K, Okuda D, Morita T, Mizuno H. Crystal structure of EMS16 in complex with the integrin alpha2-I domain. J Mol Biol. 2004;341(2):519-27. doi: 10.1016/j.jmb.2004.06.036. PubMed PMID: 15276841.

28. Jakubowski P, Calvete JJ, Eble JA, Lazarovici P, Marcinkiewicz C. Identification of inhibitors of alpha2beta1 integrin, members of C-lectin type proteins, in Echis sochureki venom. Toxicol Appl Pharmacol. 2013;269(1):34-42. doi: 10.1016/j.taap.2013.03.002. PubMed PMID: 23499869; PubMed Central PMCID: PMC3627822.

29. Momic T, Cohen G, Reich R, Arlinghaus FT, Eble JA, Marcinkiewicz C, et al. Vixapatin (VP12), a c-type lectin-protein from Vipera xantina palestinae venom: characterization as a novel anti-angiogenic compound. Toxins (Basel). 2012;4(10):862-77. doi: 10.3390/toxins4100862. PubMed PMID: 23162702; PubMed Central PMCID: PMC3496993.

30. Sarray S, Srairi N, Hatmi M, Luis J, Louzir H, Regaya I, et al. Lebecetin, a potent antiplatelet C-type lectin from Macrovipera lebetina venom. Biochim Biophys Acta. 2003;1651(1-2):30-40. PubMed PMID: 14499586.

31. Vaiyapuri S, Hutchinson EG, Ali MS, Dannoura A, Stanley RG, Harrison RA, et al. Rhinocetin, a venom-derived integrin-specific antagonist inhibits collagen-induced platelet and endothelial cell functions. J Biol Chem. 2012;287(31):26235-44. doi: 10.1074/jbc.M112.381483. PubMed PMID: 22689571; PubMed Central PMCID: PMC3406708.

32. Fukuda K, Mizuno H, Atoda H, Morita T. Crystal structure of flavocetin-A, a platelet glycoprotein Ib-binding protein, reveals a novel cyclic tetramer of C-type lectin-like heterodimers. Biochemistry. 2000;39(8):1915-23. PubMed PMID: 10684640.

33. Murakami MT, Zela SP, Gava LM, Michelan-Duarte S, Cintra AC, Arni RK. Crystal structure of the platelet activator convulxin, a disulfide-linked alpha4beta4 cyclic tetramer from the venom of Crotalus durissus terrificus. Biochem Biophys Res Commun. 2003;310(2):478-82. PubMed PMID: 14521935.

34. Morita T. Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. Toxicon. 2005;45(8):1099-114. doi: 10.1016/j.toxicon.2005.02.021. PubMed PMID: 15922777.

35. Arlinghaus FT, Eble JA. C-type lectin-like proteins from snake venoms. Toxicon. 2012;60(4):512-9. doi: 10.1016/j.toxicon.2012.03.001. PubMed PMID: 22781131.

36. Ogawa T, Chijiwa T, Oda-Ueda N, Ohno M. Molecular diversity and accelerated evolution of C-type lectin-like proteins from snake venom. Toxicon. 2005;45(1):1-14. doi: 10.1016/j.toxicon.2004.07.028. PubMed PMID: 15581677.

37. Berndt MC, Metharom P, Andrews RK. Primary haemostasis: newer insights. Haemophilia. 2014;20 Suppl 4:15-22. doi: 10.1111/hae.12427. PubMed PMID: 24762270.

38. Bryckaert M, Rosa JP, Denis CV, Lenting PJ. Of von Willebrand factor and platelets. Cell Mol Life Sci. 2015;72(2):307-26. doi: 10.1007/s00018-014-1743-8. PubMed PMID: 25297919; PubMed Central PMCID: PMC4284388.

39. Morita T. C-type lectin-related proteins from snake venoms. Curr Drug Targets Cardiovasc Haematol Disord. 2004;4(4):357-73. PubMed PMID: 15578958.

40. Bracht T, Figueiredo de Rezende F, Stetefeld J, Sorokin LM, Eble JA. Monoclonal antibodies reveal the alteration of the rhodocetin structure upon alpha2beta1 integrin binding. Biochem J. 2011;440(1):1-11. doi: 10.1042/BJ20110584. PubMed PMID: 21774787.

41. Lu C, Shimaoka M, Ferzly M, Oxvig C, Takagi J, Springer TA. An isolated, surface-expressed I domain of the integrin alphaLbeta2 is sufficient for strong adhesive function when locked in the open conformation with a disulfide bond. Proc Natl Acad Sci U S A. 2001;98(5):2387-92. doi: 10.1073/pnas.041606398. PubMed PMID: 11226249; PubMed Central PMCID: PMC30148.

42. Tuckwell DS, Smith L, Korda M, Askari JA, Santoso S, Barnes MJ, et al. Monoclonal antibodies identify residues 199-216 of the integrin alpha2 vWFA domain as a functionally important region within alpha2beta1. Biochem J. 2000;350 Pt 2:485-93. PubMed PMID: 10947963; PubMed Central PMCID: PMC1221276.

43. Mizuno H, Fujimoto Z, Koizumi M, Kano H, Atoda H, Morita T. Crystal structure of coagulation factor IX-binding protein from habu snake venom at 2.6 A: implication of central loop swapping based on deletion in the linker region. J Mol Biol. 1999;289(1):103-12. doi: 10.1006/jmbi.1999.2756. PubMed PMID: 10339409.

44. Mizuno H, Fujimoto Z, Atoda H, Morita T. Crystal structure of an anticoagulant protein in complex with the Gla domain of factor X. Proc Natl Acad Sci U S A. 2001;98(13):7230-4. doi: 10.1073/pnas.131179698. PubMed PMID: 11404471; PubMed Central PMCID: PMC34651.

45. Maita N, Nishio K, Nishimoto E, Matsui T, Shikamoto Y, Morita T, et al. Crystal structure of von Willebrand factor A1 domain complexed with snake venom, bitiscetin: insight

into glycoprotein Ibalpha binding mechanism induced by snake venom proteins. J Biol Chem. 2003;278(39):37777-81. doi: 10.1074/jbc.M305566200. PubMed PMID: 12851390.

46. Fukuda K, Doggett T, Laurenzi IJ, Liddington RC, Diacovo TG. The snake venom protein botrocetin acts as a biological brace to promote dysfunctional platelet aggregation. Nat Struct Mol Biol. 2005;12(2):152-9. doi: 10.1038/nsmb892. PubMed PMID: 15665869.

47. Horii K, Okuda D, Morita T, Mizuno H. Structural characterization of EMS16, an antagonist of collagen receptor (GPIa/IIa) from the venom of Echis multisquamatus. Biochemistry. 2003;42(43):12497-502. doi: 10.1021/bi034890h. PubMed PMID: 14580195.

48. Okuda D, Horii K, Mizuno H, Morita T. Characterization and preliminary crystallographic studies of EMS16, an antagonist of collagen receptor (GPIa/IIa) from the venom of Echis multisquamatus. J Biochem. 2003;134(1):19-23. PubMed PMID: 12944366.

49. Eble JA, Niland S, Dennes A, Schmidt-Hederich A, Bruckner P, Brunner G. Rhodocetin antagonizes stromal tumor invasion in vitro and other alpha2beta1 integrin-mediated cell functions. Matrix Biol. 2002;21(7):547-58. PubMed PMID: 12475639.

50. Navdaev A, Lochnit G, Eble JA. The rhodocetin alphabeta subunit targets GPIb and inhibits von Willebrand factor induced platelet activation. Toxicon. 2011;57(7-8):1041-8. doi: 10.1016/j.toxicon.2011.04.008. PubMed PMID: 21524659.

51. Eble JA, Beermann B, Hinz HJ, Schmidt-Hederich A. alpha 2beta 1 integrin is not recognized by rhodocytin but is the specific, high affinity target of rhodocetin, an RGD-

independent disintegrin and potent inhibitor of cell adhesion to collagen. J Biol Chem. 2001;276(15):12274-84. doi: 10.1074/jbc.M009338200. PubMed PMID: 11121411.

52. Fontana A, Scoffone E. [40] Sulfenyl halides as modifying reagents for polypeptides and proteins. Methods Enzymol. 1972;25:482-94. doi: 10.1016/S0076-6879(72)25044-3. PubMed PMID: 23014429.

53. Eble JA, Tuckwell DS. The alpha2beta1 integrin inhibitor rhodocetin binds to the A-domain of the integrin alpha2 subunit proximal to the collagen-binding site. Biochem J. 2003;376(Pt 1):77-85. doi: 10.1042/BJ20030373. PubMed PMID: 12871211; PubMed Central PMCID: PMC1223739.

54. Leslie AGW. MOSFLM users guide. MRC-LMB, Cambridge. 1994.

55. CCP4. Collaborative Computing Project No. 4, The CCP4 suite: programs for protein crystallography. Acta Cryst D50. 1994:760-3.

56. Ozbek S, Muller JF, Figgemeier E, Stetefeld J. Favourable mediation of crystal contacts by cocoamidopropylbetaine (CAPB). Acta Crystallogr D Biol Crystallogr. 2005;61(Pt 4):477-80. doi: 10.1107/S0907444905001204. PubMed PMID: 15805603.

57. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997;53(Pt 3):240-55. doi: 10.1107/S0907444996012255. PubMed PMID: 15299926.

58. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501. doi: 10.1107/S0907444910007493. PubMed PMID: 20383002; PubMed Central PMCID: PMCPMC2852313.

59. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, et al. The Phenix software for automated determination of macromolecular structures. Methods. 2011;55(1):94-106. doi: 10.1016/j.ymeth.2011.07.005. PubMed PMID: 21821126; PubMed Central PMCID: PMCPMC3193589.

Chapter 5: Archaea S-layer nanotube from a “black smoker” in complex with cyclo-octasulfur (S_8) rings

Matthew McDougall[1,2], Olga Francisco[1,2], Candice Viddal[5], Roy Roshko[6], Markus Meier[1] and Jörg Stetefeld[1,2,3,4]

[1] Department of Chemistry, University of Manitoba

[2] Center for Oil and Gas Research and Development (COGRAD)

[3] Department of Biochemistry and Medical Genetics, University of Manitoba

[4] Department of Human Anatomy and Cell Science, University of Manitoba

[5]Canadian Mennonite University, 500 Shaftesbury Blvd, Winnipeg, Manitoba, Canada

[6]Department of Physics and Astronomy, University of Manitoba

First published in *Proteins: Structure, Function, and Bioinformatics*, **85,** 2209-2216 (2017)

DOI 10.1002/prot.25385

Preface

This article, published in *Proteins: Structure, Function, and Bioinformatics,* explores the ability of an extracellular protein of a unicellular organism to sequester key metabolic substrates from a harsh and unforgiving environment. The low solubility and spectroscopic activity of elemental sulfur, as well as the high temperature required for the binding event, present considerable challenges in examining the energetics of the incorporation experimentally. Instead, by combining structural data with computational analysis we could demonstrate that RHCC is not only capable of bind elemental sulfur, but is likely to do so in its native environment.

Contributions of Authors

All computational work was completed by Candice Viddal and Roy Roshko of the Canadian Mennonite University and the University of Manitoba respectively. Matthew McDougall (University of Manitoba) and Olga Francisco (Center for Oil and Gas Research and Development) expressed and purified RHCC. Matthew McDougall obtained and crystallized the RHCC-S_8 complex and was supported in data collection and structure solution by Markus Meier and Jörg Stetefeld (University of Manitoba). The manuscript was prepared by Matthew McDougall, Roy Roshko, and Jörg Stetefeld.

Abstract

Elemental sulfur exists primarily as an S_8^0 ring and serves as terminal electron acceptor for a variety of Sulfur-fermenting bacteria. The large quantities of inorganic sulfur that black smoker vents emit makes archaea living there an exciting research tool to advance our knowledge of sulfur respiration. Here we use a hybrid method approach to demonstrate that the proteinaceous cavities of the S-layer nanotube of the hyperthermophilic archaeon *Staphylothermus marinus* act as a storage reservoir for cyclo-octasulfur (S_8). Fully atomistic molecular dynamics simulations were performed and the method of Multi-Configurational Thermodynamic Integration was employed to compute the absolute free energies for transferring nine water molecules and a ring of eight sulfur atoms (S_8) individually from an aqueous bath into one of the larger hydrophobic cavities of a fragment of archaeal tetrabrachion. The calculations show that the sulfur ring is energetically favoured in the cavity over the water cluster.

Introduction

Sulfur is an essential element of life. The use of elemental sulfur (S^0) as the terminal electron acceptor (TEA) is a common theme in hyperthermophilic archaea, progressing through a number of unique and not fully understood metabolic pathways [1-3]. Many of these microorganisms were discovered in volcanic environments (such as submarine hydrothermal systems, hot springs or terrestrial solfataric fields), which produce elemental sulfur as a key substrate in high concentrations at different temperature regimes [4]. The most abundant form of elemental sulfur is the ring-like S_8 crown. However, its extreme hydrophobicity and its low solubility in water (only 5μg/L at 25°C) makes it poorly suited as a substrate for bacterial sulfur respiration.

Since their discovery, hyperthermophilic archaea have been under intense scrutiny due to their unique adaptations to the extreme environments they thrive in [5]. It is suggested that life has hyperthermophilic origins, and thus the study of these organisms provides insight into the biology of early life forms. Most known hyperthermophiles are dependent on S^0 reduction and use either H_2 or organic compounds as electron donors [6]. *Staphylothermus marinus* (*S. marinus*), isolated from the periphery of black smoker vents, is a strict anaerobic sulfur reducing archaeon that utilizes peptides as a carbon source and grows at temperatures up to 98°C [7]. S^0 is strictly required as a TEA for its growth; it is reduced to H_2S, and does not serve as physical support [8]. *S. marinus* has a proteinaceous S-layer consisting of elongated tetrabrachion units arranged in a 2D-network that coats the archaeon [9, 10]. Freeze-etching cuts of reconstituted S-layers from *S.marinus* showed a strong red coloring after addition of ruthenium, suggesting a major storage space for elemental sulfur [2].

A 24 kDa segment of this tetrabrachion forms the tetrameric right handed coiled-coil structure (RHCC). RHCC is a cysteine-free protein segment with a pI value of 4.3 and four large cavities in the interior of the tetrameric channel [11]. In its natural environment, RHCC is associated with two subtilisin-like proteases known as STABLE [12].

To date, it is not known how elemental sulfur can be stored in a proteineacous environment and made accessible for biochemical reactions. Here, we present the 1.9Å crystal structure of RHCC in complex with S_8 ring-like structures inside the large cavities of the nanotube. To our knowledge, this is the first study describing the structure of protein cavities complexed with S_8 crowns in a cysteine-free environment. Molecular dynamic simulations presented here reveal that the incorporation of S_8 ring systems is energetically favorable and enhances the stability over clathrate-like water structures. To probe the mechanism of entry to the cavities, we performed thermal Dynamic Light Scattering experiments, revealing that an intact tetrameric channel system is flexible enough to allow access of large space filling ring structures.

Results

Biophysical and structural elucidation of the $RHCC - S_8$ complex

Elemental sulfur was introduced into the cavities by heating RHCC to 80ºC in the presence of sublimed sulfur for at least 96 hours, which mimics a black smoker environment in a laboratory setting (**Figure 5-1**). The DLS melting curves showed that RHCC-NT is monodisperse and homogenous in solution under these extreme conditions and does not change its tetrameric assembly nor its hydrodynamic properties (**Figure 5-2**). CD measurements under different temperature regimes confirms the maintenance of α-helicity (**Appendix 3, Figure III-1**). To unravel the molecular features of the S_8-uptake, we determined the crystal structure of $RHCC - S_8$ at 1.9Å resolution (λ=1.54) in comparison to a long-wavelength S-SAD experiment at λ=1.77Å (**see Table 5-1 and Appendix 3, Figure III-2**).

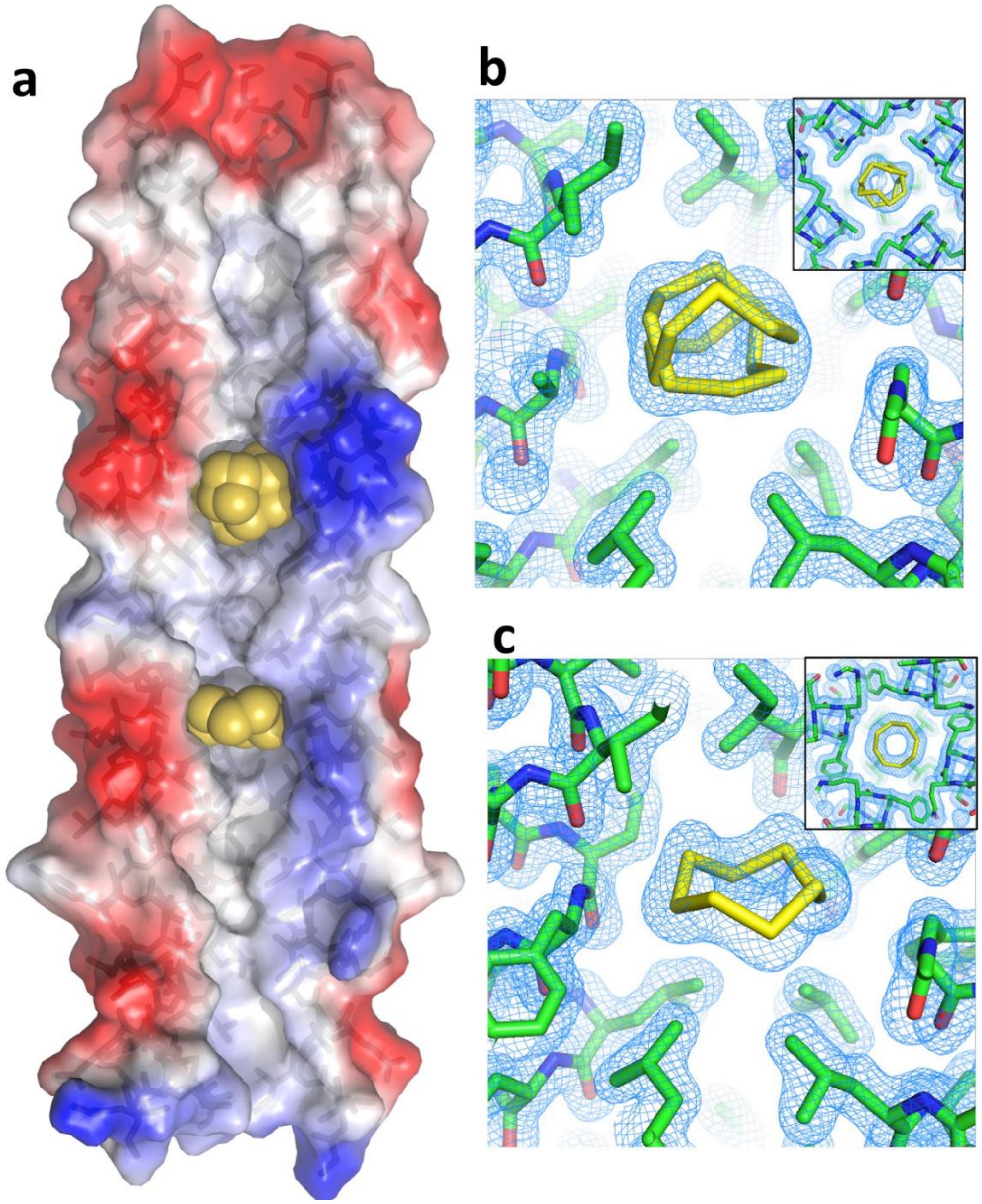


Figure 5-1. RHCC in complex with cyclo-octasulfur S_8 crowns.

(**a**) Cross section of the tetrameric RHCC channel. The molecular surface of the protein component is shown in white, with positively and negatively charged side chains colored red and blue, respectively. The individual S_8-entities are highlighted as yellow spheres. (**b**) and (**c**) The 1.9Å 2Fo-Fc map (1.2σ contour level) are shown for Cavity 3 (**b**) and Cavity 2 (**c**) perpendicular to the helical axis of the tetrameric channel. The S_8-crowns are highlighted in yellow. The inset shows the orthogonal view, along the tetrameric channel of RHCC.

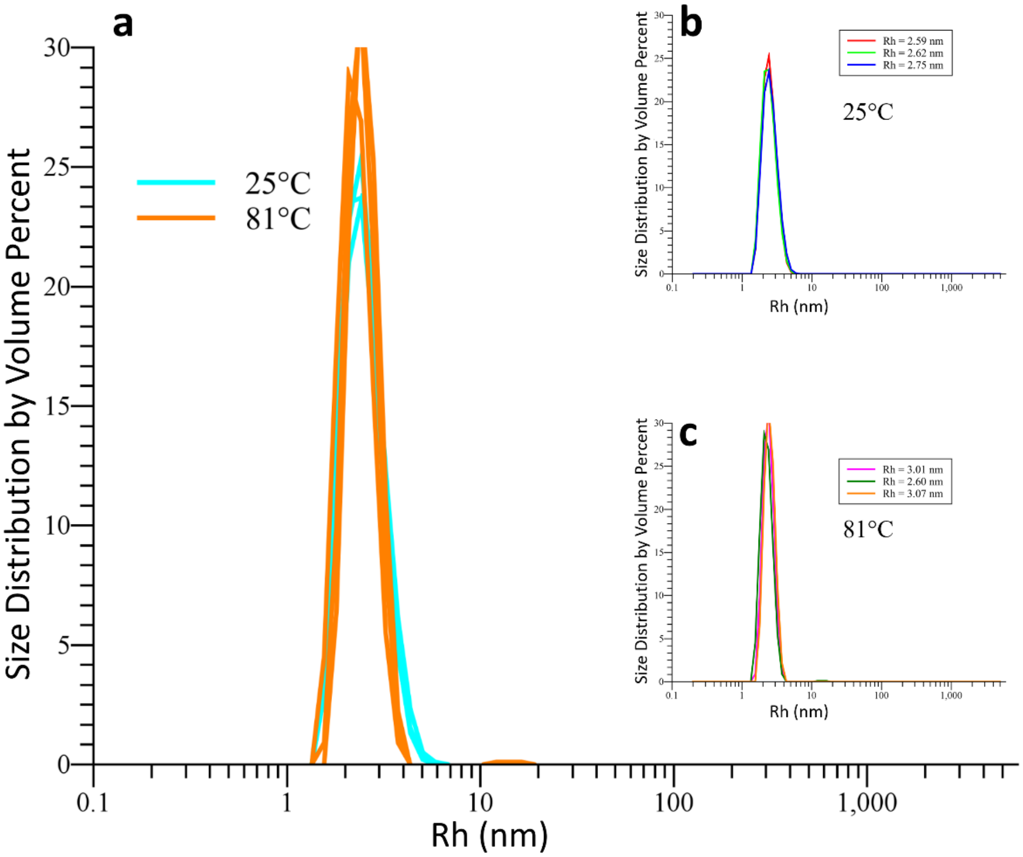


Figure 5-2. Thermostability of the RHCC tetramer.

(**a**), Overlay of Dynamic Light Scattering curves which were collected in triplets at 25ºC and 81ºC. The hydrodynamic radius (R_h) at 25ºC (**b**) is 2.65 ± 0.09nm, in comparison to 2.89 ± 0.26 at 81ºC (**c**). RHCC exhibits minimal change in hydrodynamic radius with increased temperature, which is consistent with a correctly folded tetramer during the uptake of S_8.

In the RHCC-nanotube, the cavities are connected by a narrow and continuous channel that is lined exclusively with aliphatic sidechains of the hydrophobic core residues in the a and h position of the undecad repeat [11, 13]. The diameter of the channel varies between 2Å and 8Å. The 1.9Å crystal structure of $RHCC - S_8$ revealed that both larger cavities (named here cavity 2 and 3) of the RHCC nanotube are occupied with ring-shaped elemental S_8 sulfur (**Figure 5-1**).

Cavity 2, which is confined by knobs-into-holes residues Ile19 (*a*) and Leu26 (*h*), has a volume of 380 $Å^3$ and is gated by a complex ionic network surrounding the DDRYES motif[11]. The $S_8 -$ crown occupies a total volume of 331$Å^3$ and is only involved in van der Waals contacts along the Cα-vector between Arg22 and Tyr23, as well as β-branched forks of Ile19 and Leu26 (**Figure 5-1b and Appendix 3, Figure III-2**). The electron-sink of the cyclo-octasulfur is perpendicular to the axis of the coiled coil dipole. The sulfur $S_8 -$ crown has geometric parameters similar to that of crystalline $\alpha - S_8$, forming a crown-shaped ring system with bond lengths of 2.07 Å and angles of 108.2° (see also **Table 5-2**). In cavity 3, which is defined by Ile30 (*a*) and Leu37 (*h*), the flexible sulfur ring adopts multiple conformations that occupy the entire 354$Å^3$ (**Figure 5-1c and Appendix 3, Figure III-1**). Conformation A has a relative geometry similar to the twisted-ring S_8, whereas Conformation B adopts the *exo-endo* $- S_8$ arrangement. Unrestrained crystallographic refinement protocols reveal an equi-molar occupancy where both $S_8 -$ ring systems tumble around the co-planar axis of the cavity.

Taken together, these results illustrate that both large-size cavities of the RHCC-nanotube are able to uptake elemental sulfur. The sulfur rings have near ideal geometry, with bond angles between 99.2º and 117.8º and bond distances with an average of 2.07Å. These compare well with the ideal values of 108° and 2.06Å, respectively [14]. The breathing effect present during the uptake of vitamin D3 in COMPcc is missing [13] as the RMSD in Cα-positions is less than 0.2Å and the overall B-factor distribution is comparable to $apo - RHCC$ (PDB:1FE6).

Stability of RHCC-S_8 vs apo-RHCC

In the $apo - RHCC$ structure of archaeal tetrabrachion, cavity 2 is occupied by nine ordered water molecules whose positions correspond closely to those of the sulfur atoms in the $RHCC - S_8$ structure (**Figure 5-3**). Remarkably, the clathrate-like nine-water cluster reveals a non-planar ring conformation, which is reminiscent of the geometry of the $\alpha - S_8$ crown although it lacks the structural rigidity of the sulfur crown. Fully atomistic molecular dynamics simulations were performed on both the $apo - RHCC$ and the $RHCC - S_8$ complexes and the method of Multi-Configurational Thermodynamic Integration (MCTI) [15-22] was employed to calculate, individually, the absolute free energies of transfer of nine water molecules and a crown of eight sulfur atoms S_8 from an aqueous bath into hydrophobic cavity 2. The standard state transfer free energy for sulfur was determined to be $\Delta G^0_{S_8} = -55.5 \pm 2.9\ kJ/mol$ while the standard state transfer free energy for water was $\Delta G^0_{[H_2O]_9} = -5.2 \pm 4.6\ kJ/mol$, showing that the sulfur crown is preferred over water by $\Delta\left(\Delta G^0_{transfer}\right) = \Delta G^0_{S_8} - \Delta G^0_{[H_2O]_9} = -50.3 \pm 5.4\ kJ/mol$ (See **Table 5-3**). It is expected that the effect is cumulative over all the cavities.

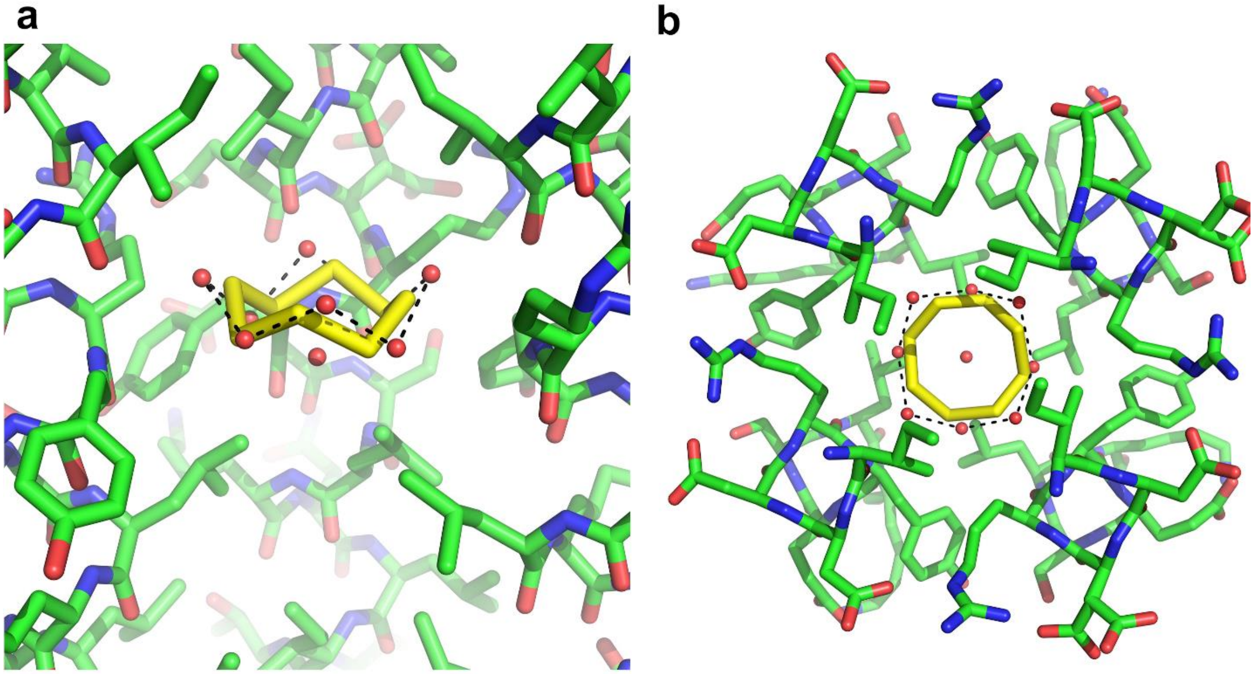


Figure 5-3. Comparison of the nine-water cluster with the S_8 crown in cavity 2.

Structural models of the nine-water cluster in superimposition with the S_8 crown in cavity 2 in a perpendicular (**a**) and parallel (**b**) view to the helical axis of RHCC. The 8 waters in the ring conformation are not planar; instead, they form a crown shape reminiscent of α-S_8. Indeed, superposition of α-S_8 onto these ordered waters results in an almost perfect overlap of the Van der Waals radii.

Discussion

In the apo-RHCC structure, all cavities of archaeal tetrabrachion were filled with water molecules (**Figure 5-3**). Yin *et al.* [23] examined the thermodynamic and structural properties of water in the large cavities of archaeal tetrabrachion as a function of temperature and solvent properties. By calculating the free energy of transfer of water from the bulk into the interior, Yin *et al* found that occupied cavities are marginally favourable at 298K and 1 atm for occupancies $6 \leq N \leq 9$, with a minimum in the free energy of transfer at an occupancy of $N = 8$, consistent with the x-ray studies which reveal an occupancy of $N = 9$. As the most common allotrope of sulfur is S_8 we were led to inquire whether it may be thermodynamically more favorable for S_8 to occupy the cavity instead of water, given that *S. marinus* is found around black smoker vents, an environment rich in sulfur. The molecular dynamics simulations and free energy calculations performed in this investigation show that it seems likely that RHCC in its native environment does not contain water within the cavities, but instead contains S_8 crowns.

Many archaea utilize sulfur compounds as electron donors or acceptors for energy production. However, it is currently not known, how circularized S_8 sulfur is carried in these microorganisms. The cavities of RHCC are hydrophobic in nature and are too large in size to be left empty in the properly folded protein [13, 24, 25]. The ability for tetrabrachion to store sulfur compounds in large quantities is particularly striking because it has been shown that *S. marinus* is a strict anaerobic hyperthermophilic archeon that uses S^0 as a TEA [3].

Our findings raise the question of why the elemental sulfur is stored inside the cavities and how it is made accessible as a TEA. We suggest that the cavities are acting as a storage reservoir, which provides S_8 in required quantities and sequesters S_8 from nanocrystals in the environment. Boyd and Druschel have shown that large S^0 nanoparticles hamper growth rates of sulfur-reducing microorganisms [26]; sequestering individual sulfur molecules is likely critical for cellular respiration.

A remaining question concerns the availability of S_8-crowns inside the hydrophobic cavities as a TEA. *S. marinus* is a S^0-reducing archaea with sulfur reductase activities comparable to those in other sulfur-reducing prokaryotes [8]. The authors propose a mechanism whereby reducing equivalents are generated by the oxidation of amino acids from STABLE-catalyzed peptidolysis, which then reduce S^0 and protons to H_2S and H_2, respectively. Remarkably, the RHCC-STABLE complex is one of the most stable noncovalent protein-protein interactions described thus far, and STABLE is the only known enzyme that is a stoichiometric component of an archaeal S-layer. However, elucidation of the molecular mechanisms of S^0-dependent metabolism requires further experimentation.

Tables

Table 5-1. Crystallographic table (Sulfur-native/ S-SAD)[a]

Data Collection	RHCC-S	RHCC-S ano
λ (Å)	1.54192	1.77122
Space Group	P31 2 1	P31 2 1
Cell dimensions		
a, b, c (Å)	110.52 110.52 71.04	110.71 110.71 70.92
α, β, γ (°)	90.00 90.00 120.00	90.00 90.00 120.00
No. reflections	194980 (23775)	966799 (98939)
Resolution (Å)	25-75 – 1.90 (2.00 – 1.90)	19.84 - 2.40 (2.49 – 2.40)
[b]R_{merge}	0.121 (0.826)	0.101 (0.594)
$I/\sigma I$	6.2 (1.70)	40.7 (10.0)
Completeness (%)	98.9 (91.6)	99.8 (100.0)
Multiplicity	4.9 (4.2)	48.5 (48.1)
Anomalous Completeness (%)	-	99.9 (100.0)
Anomalous Multiplicity	-	25.0 (24.4)
Refinement		
[c]R_{work}/R_{free}	0.182 / 0.207	
No. atoms		
Protein	3365	
Ligand/Ion	39	
Water	328	
B-factor (Å^2) Protein/Water/ S_8 ligands	33.90 31.7/44/ 47.1/50.9/57.7	
R.m.s. deviations		
Bond lengths (Å)	0.0115	
Bond angles (°)	1.106	

[a]Statistics of the highest resolution shell are shown in parenthesis. [b]$R_{merge} = \sum_{hkl}\sum_{i=1}^{N}\left|I_{hkl,i} - \bar{I}_{hkl}\right| / \sum_{hkl}\sum_{i=1}^{N} I_{hkl,i}$ [c]$R_{factor} = \sum_{hkl}\left|F_{hkl}^{obs} - F_{hkl}^{calc}\right| / \sum_{hkl} F_{hkl}^{obs}$ The R_{free} was calculated by randomly omitting 10% of observed reflections from refinement.

Table 5-2. Geometry of individual S_8 ring systems.

	Cavity 2[$] crown-$S_8$	Cavity 3 Twisted-ring S_8	Cavity 3 Exo-endo-S_8
Bond Angle (°)			
8-1-2	112.7	105.4	110.7
1-2-3	106.4	108.0	110.1
2-3-4	108.5	99.2	114.6
3-4-5	107.6	106.7	111.3
4-5-6	104.9	106.3	107.1
5-6-7	106.9	110.9	112.2
6-7-8	109.5	113.2	112.1
7-8-1	108.8	117.8	107.3
Average	108.2	108.4	110.7
Bond length (Å)			
1-2	2.00	2.05	2.08
2-3	2.12	2.08	2.07
3-4	2.06	2.04	2.02
4-5	2.11	2.07	2.08
5-6	2.05	2.09	2.09
6-7	2.10	2.11	2.10
7-8	2.06	2.03	2.04
8-1	2.09	2.06	2.11
Average	2.07	2.07	2.07

[$] Deviations in bond length from crystalline S_8 (2.037Å) may be due to the data resolution, or to the fact molecules in the cavities behave more closely to those in the gas phase than in the crystalline phase[24]. Bond lengths in the gas phase are predicted to be 2.08 Å. The nomenclature for individual S_8-ring systems is according to Wong *et al.*[14]

Table 5-3. Calculated Contributions to the Absolute Free Energy of Transfer (kJ/mol).

Ligand	**Protein →Gas**			**Solvent →Gas**		**Absolute Transfer Free Energy**
	ΔG_1	ΔG_2	ΔG_3	ΔG_4	ΔG_5	$\Delta G^0_{transfer}$
S_8	$+5.8$	$+40.9 \pm 2.8$	-2.2	$+4.4$	-15.4 ± 0.8	-55.5 ± 2.9
$[H_2O]_9$	$+6.5$	$+251.6 \pm 3.2$	2.4	-4.1	$+259.4 \pm 3.3$	-5.2 ± 4.6

Methods

Expression and Purification of RHCC

The RHCC-His_6 gene, which contains a thrombin cleavage site, was inserted into a pET15b vector and transformed into an *E. coli* strain BL21/DE3. Cells were inoculated into 50 mL of LB containing 0.1 mg/mL ampicillin. This pre-culture was allowed to grow for 12-16 hours at 37°C, then added to 1.8 L of LB 0.1 mg/mL ampicillin and grown to an OD_{600} 0.6 before they were induced with 0.8 mM IPTG. This culture was grown for 3 hours before the cells are pelleted and resuspended in a 20 mM phosphate buffer at pH 7, 2 mM imidazole, 6M guanidine hydrochloride. The cells are then lysed, sonicated, and centrifuged to remove cellular debris. RHCC-His_6 is then purified on a Talon cobalt affinity column followed by cleavage of the His_6 tag with thrombin. Protein quality/monodispersity is then assessed with FPLC and DLS.

DLS Experiments

The thermostability of RHCC was assessed by means of the Nano-S Dynamic Light Scattering system (Malvern Instruments Ltd, Malvern, UK), which employs a 633 nm laser and a fixed scattering angle of 173°. A 150 µL sample of 2 mg/mL RHCC in 20 mM tris pH 8.0, I= 154 mM (NaCl) in a 12 µL quartz cuvette was allowed to equilibrate at 25°C for 5 minutes prior to measurement in triplicate by the DLS. The same sample was then heated to 81°C and again allowed to equilibrate for 5 minutes prior to data collection by DLS.

CD Experiments

CD spectra were obtained using a J-810 spectropolarimeter (Jasco Inc.) and a 0.1 cm quartz cell (Hellma). Spectra were collected on 0.2mg/mL RHCC in 12mM phosphate, pH 7.0, 200mM NaF at 20°C and 81°C, with 30 minutes of thermal equilibration. All data were measured in triplicate with a scan speed of 7.5 nm/min.

Crystallization, Data processing and Structure refinement

15 mg of sublimed S_8 was added to 370 µl of 8.15mg/mL RHCC in a 10 mM bicine buffer (pH=8, I= 154 mM NaCl) and was incubated for 4 days at 80ºC. Excess sulfur was removed by centrifugation. Monodispersity of the resulting RHCC-S sample was confirmed by DLS prior to

concentration to 11 mg/mL. RHCC-S crystals were grown by hanging-drop vapour diffusion at 293 K by mixing 2 µL of protein solution with 2 µL reservoir solution containing 1.6M ammonium sulfate and 100 mM Tris pH 8.0. Crystals appeared after 1 week and were soaked in mother liquor containing 20% glycerol for 5-10 min before flash freezing in liquid nitrogen. Diffraction data was collected on a Rigaku MM-007HF (λ=1.54178Å) at 100 K. The dataset was indexed, integrated and scaled with MOSFLM [27] and the CCP4-package [28]. The phases were determined by rigid body refinement using a previously solved RHCC structure (PDB code 1YBK) in Refmac [29]. The model was built and refined without NCS-restraints using Coot and refined using the Phenix software package [30, 31]. Anomalous data was collected at λ=1.77123Å at the 08B1-1 beamline at the Canadian Light Source. Data was indexed, integrated, and scaled with XDS [32] and the CCP4-package. The difference map was calculated using experimental amplitudes (F_{ano}) and calculated phases (φ_{calc}) from apo-RHCC (PDB: 1FE6). The final coordinates and structure factor amplitudes were deposited (PDB: 5JR5).

Molecular Dynamic simulations

Molecular dynamics simulations of the apo-RHCC and RHCC-S_8 complexes were performed with the GROMACS molecular dynamics simulation package [33] using the Gromos 43a2 force field with an SPC water model. The crystal structure of the apo-RHCC complex, including the measured coordinates of the cluster of nine water molecules in cavity 2, was previously determined by Stetefeld *et al.* to a resolution of 1.8Å and is available in the RCSB Protein Data Bank (PDB: 1FE6)[11]. The coordinates of the S_8 crown were obtained from Eckert *et al.* [34].

The topology of the sulfur crown was generated by PRODRG [35]. This yielded charges of zero on all sulfur atoms (comparable to those farthest away from nitrogen in heptasulfur imide (S7NH) calculated using DFT, PDFT/MM, and KRMM [36]), Lennard-Jones parameters $C_{12} = 1.31 \times 10^{-5}\ kJ\ nm^{12}\ mol^{-1}$ and $C_6 = 0.00998\ kJ\ nm^6\ mol^{-1}$, equilibrium S-S bond lengths $b_0 = 0.204\ nm$, equilibrium S-S-S bond angles $\theta_0 = 104.0^o$ and force constants $k_b = 5.03 \times 10^6\ kJ\ mol^{-1} nm^{-2}$ and $k_\theta = 490.0\ kJ\ mol^{-1} rad^{-2}$. The usual combination rules were employed, specifically, $C_{ij}^{(6)} = \left(C_{ii}^{(6)} C_{jj}^{(6)}\right)^{1/2}$ and $C_{ij}^{(12)} = \left(C_{ii}^{(12)} C_{jj}^{(12)}\right)^{1/2}$.

The RHCC complexes were solvated in a rectangular box with dimensions 5.3 nm x 5.3 nm x 10.4 nm, to which 8681 SPC water molecules and 16 charge neutralizing Na^+ ions were added. The closest distance between any point on the RHCC complex and the surface of the simulation box was 1.2 nm. Long-range electrostatic interactions were treated with the particle mesh Ewald method. The cut-off radius for all non-bonded interactions was 1.0 nm and the neighbour list for all non-bonded interactions was updated every 10 time steps. The measured structure was initially energy minimized to 1000 kJ/mol using the method of steepest descent and then minimized further to 600, 400, 200, 100, 90 and 80 kJ/mole before convergence was achieved, typically around 50-60 kJ/mol for S_8 and 30-40 kJ/mol for the group of nine waters.

The energy minimized system was first heated gradually from 0 K to 50K over a 20 ps time period while position-restraining the non-hydrogen atoms, using velocities from a Maxwell-Boltzmann distribution. Equilibration of all atoms was achieved within 3ps. Subsequent heatings to 150K and 300K were also performed over a 20 ps time period with the initial velocities at each new temperature taken from the final frame of the equilibrated trajectory at the previous temperature. During the production run, the entire system (RHCC tetrabrachion plus cavity contents plus solvent bath) was maintained at a temperature of 300K and a pressure of 1 atm using Berendsen thermostats and barostats, respectively, with time constants of 0.1 ps. Production runs were performed for 2 ns in time steps of 1 fs with temperature and pressure scaling. Simulations of the solvated protein containing the ligand (S_8 or $(H_2O]_9$) and of the solvated ligand alone were performed separately in periodic boxes of the same dimensions.

Calculations of the absolute standard state transfer free energy for the S_8 sulfur crown and the nine water molecules from the bulk into cavity 2 of RHCC tetrabrachion were based on the two thermodynamic paths shown in **Appendix 3, Figure III-3**. **(**It was assumed that the water molecules enter the cavity individually and, once inside, assemble into a cluster under the influence of their mutual hydrogen bonding forces). The upper path shows the removal of a ligand L from a bound, fully interacting, solvated ligand-protein complex (LP_{sol}) to the gas phase. The free energy difference ΔG_1 describes the effect of restraining the translational motion of the fully interacting ligand by turning on a flat bottomed harmonic well (FBHW) [19] of the form $U_{FBHW}(r < r_0) = 0$; $U_{FBHW}(r > r_0) = k(r - r_0)^2$ with $k = 1000\ kJ\ mol^{-1}\ nm^{-2}, r_0 = 0.55\ nm$, ΔG_2 describes the transformation of the ligand into a noninteracting dummy particle in the gas phase by gradually turning off the ligand interactions and ΔG_3 converts the free energy of the restrained ligand to the standard state volume V_0. (No conformational restraints or angle restraints were applied as the tetrabrachion cavities act merely as storage receptacles and do not contain active site binding receptors). The lower path shows the same sequence for the removal of a ligand L from the solvent to the gas phase in the absence of RHCC tetrabrachion. No further standard state correction was necessary here as the non-interacting ligand states in the two paths are identical. The point symmetry group D4d of the S_8 ligand corresponds to a rotational symmetry number σ=16 and, in principle, the existence of multiple, symmetrically equivalent states would lead to an additional contribution $\Delta G_{symm} = - RT\ ln16\ =\ -6.9\ kJ\ mol^{-1}$ to the free energy. However, as no body restraints have been applied that would break the rotational symmetry of the molecule by artificially restricting the molecule to a particular orientation, no such correction has been included in the current analysis. The absolute standard state free energy for the transfer of the ligand from the solvent bath to the cavity is then given by

Equation 5-1

$$\Delta G^0_{transfer} = - \Delta G_1 - \Delta G_2 - \Delta G_3 + \Delta G_4 + \Delta G_5\ .$$

The free energy differences ΔG_1 and ΔG_4 which describe the effect of applying the FBHW restraint were computed numerically using thermodynamic perturbation [19]. The terms ΔG_2 and ΔG_5 which describe the transfer of the fully interacting ligand to the non-interacting gas phase were computed using the method of Multi-Configurational Thermodynamic Integration (MCTI) [15-22] with a single topology format. The nonbonded parameters were scaled with a coupling parameter λ which varied linearly from λ=0 (ligand fully interacting) to λ=1 (ligand non-interacting) with both the bond lengths and bond angles of the dummy atoms fixed. The standard state correction ΔG_3 was calculated analytically using $\Delta G_3 = -RTln(V_{MD}/V_0)$ where V_{MD} is the FBHW simulation volume and V_0 is the standard state volume with $V_0 = 1.660\ nm^3$ when the ligand is the sulfur crown and $V_0 = n \times 0.030\ nm^3$ when the ligand is a group of n water molecules [19].

The energy minimization and thermal equilibration procedures were repeated for each value of λ before generating production runs for 2 ns. Structures were saved for analysis every 0.5 ps during the production runs and the resulting set of 4000 structures was used to determine $\langle dG/d\lambda \rangle$ for each value of λ. Each $\langle dG/d\lambda \rangle$ plot was generated using fifteen values of λ. The mean value $\langle dG/d\lambda \rangle$ was calculated separately for both the electrostatic and the van der Waals non-bonded contributions to the free energy using g_analyze of the Gromacs MD package and the errors were computed by block averaging. The statistical error in $\langle dG/d\lambda \rangle$ was determined from the autocorrelation of the data over the 2ns trajectory.

Soft-core interactions were incorporated into all simulations presented in this paper in order to remove the discontinuity in the non-bonded parameters as the interactions are turned off and to allow for proper convergence. A soft-core parameter α = 0.1 was found to produce the smoothest free energy curves and the best convergence.

The sampling of configuration space is believed to be sufficiently extensive during end-state removals as confirmed by the low statistical error in $\langle dG/d\lambda \rangle$ in the non-interacting limit $\lambda \rightarrow 1$, which is typically between 0.5% and 3%. For all λ steps, the values of $\langle dG/d\lambda \rangle$ converged to a time-independent constant within the first 1ns of each 2ns trajectory and the derivative $dG/d\lambda$ varied smoothly with λ, both of which are indicative of convergence.

Table 5-3 summarizes the calculated values of all five contributions to the transfer free energy for both the sulfur crown and the water cluster. In order to achieve agreement with the measured solvation free energy of S_8, it was necessary to modify the non-bonded Lennard-Jones interaction parameters between the sulfur atoms and the water oxygens [37, 38]. An increase in the repulsive parameter C_{12} for $S - OW$ by a factor of 1.4 yielded a calculated solvation free energy for S_8 of $\Delta G^0_{solvation}(S_8) = -\Delta G_4 - \Delta G_5 = +11.0\ kJ\ mol^{-1}$ which is consistent with the experimental estimate of $\Delta G^0_{solvation}(S_8) = +15.5\ kJ\ mol^{-1}$ (standard states 1M in solution and 1 bar in the gas phase, solid activity =1) based on the measured solid state solubility of $1.9\ \times\ 10^{-8}\ mol/kg$ [39] and solid state vapour pressure of 10^{-5} bar at 39°C [40]. For purposes of consistency, the same factor of 1.4 was applied to the interactions between the ring-sulfurs and all atoms in the cavity walls $(S - O, S - N, S - C)$. Similarly, in order to replicate the rigid 9-water clathrate cluster observed experimentally in the cavity, it was necessary to modify the non-bonded Lennard-Jones interaction parameters between the water oxygens and the aliphatic carbon atoms in the cavity walls [41]. An increase in the attractive parameter C_6 for $OW - C$ by a factor of 1.7 yielded a free energy of transfer for $N = 9$ water molecules from the bulk into cavity 2 of tetrabrachion of $\Delta G^0_{[H_2O]_9} = -\ \Delta G_1 -\ \Delta G_2 -\ \Delta G_3 +\ \Delta G_4 +\ \Delta G_5 =\ -5.2 \pm 4.6\ kJ/mol$, which compares favourably with the value $\Delta G^0_{[H_2O]_9} = -8.19 \pm 0.81\ kJ/mol$ obtained by Yin *et al.* [23] using AMBER 6.0 with the parm94 force field and a different method of analysis. Both analyses indicate that the water cluster is marginally stable in cavity 2.

The calculations show that the water cluster is energetically preferred over the sulfur ring in both the aqueous solvent and the tetrabrachion cavity, $\Delta G(gas \rightarrow cavity) = [-\Delta G_1 - \Delta G_2 - \Delta G_3 = -260.5\ kJ/mol]_{[H_2O]_9} < [-\Delta G_1 - \Delta G_2 - \Delta G_3 = -44.5\ kJ/mol]_{S_8}$ and $\Delta G(gas \rightarrow solvent) = [-\Delta G_4 - \Delta G_5 = -255.3\ kJ/mol]_{[H_2O]_9} < [-\Delta G_4 - \Delta G_5 = +11.0\ kJ/mol]_{S_8}$. However, the net effect is a preference for the sulfur crown over the water cluster in the cavity due the non-polar nature of S_8 which results in a low aqueous solubility and a more favourable transfer free energy from the gas phase into the apolar environment of the cavity.

References

[1] Kasting, J. F., Zahnle, K. J., Pinto, J. P., and Young, A. T. (1989) Sulfur, ultraviolet radiation, and the early evolution of life, *Orig Life Evol Biosph 19,* 95-108.

[2] Hedderich, R., Klimmek, O., Kroger, A., Dirmeier, R., Keller, M., and Stetter, K. O. (1999) Anaerobic respiration with elemental sulfur and with disulfides, *FEMS Microbiology Reviews 22,* 353-381.

[3] Liu, Y., Beer, L. L., and Whitman, W. B. (2012) Sulfur metabolism in archaea reveals novel processes, *Environ Microbiol 14,* 2632-2644.

[4] Stetter, K. O. (2006) History of discovery of the first hyperthermophiles, *Extremophiles : life under extreme conditions 10,* 357-362.

[5] Stetter, K. O. (2006) Hyperthermophiles in the history of life, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences 361*, 1837-1842; discussion 1842-1833.

[6] Schicho, R. N., Ma, K., Adams, M. W., and Kelly, R. M. (1993) Bioenergetics of sulfur reduction in the hyperthermophilic archaeon Pyrococcus furiosus, *Journal of bacteriology 175,* 1823-1830.

[7] Fiala, G., Stetter, K. O., Jannasch, H. W., Langworthy, T. A., and Madon, J. (1986) Staphylothermus marinus sp. nov. represents a noverl genus of extremely thermophillic submarine heterotrophic archaeabacteria growing up to 98C, *Syst Appl Microbiol 8*, 106-113.

[8] Hao, X., and Ma, K. (2003) Minimal sulfur requirement for growth and sulfur-dependent metabolism of the hyperthermophilic archaeon Staphylothermus marinus, *Archaea 1,* 191-197.

[9] Peters, J., Baumeister, W., and Lupas, A. (1996) Hyperthermostable surface layer protein tetrabrachion from the archaebacterium Staphylothermus marinus: evidence for the presence of a right-handed coiled coil derived from the primary structure, *Journal of molecular biology 257,* 1031-1041.

[10] Peters, J., Nitsch, M., Kuhlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., Engelhardt, H., Pfander, J. P., Muller, S., Goldie, K., and et al. (1995) Tetrabrachion: a filamentous archaebacterial surface protein assembly of unusual structure and extreme stability, *Journal of molecular biology 245*, 385-401.

[11] Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., and Kammerer, R. A. (2000) Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer, *Nature structural biology 7*, 772-776.

[12] Mayr, J., Lupas, A., Kellermann, J., Eckerskorn, C., Baumeister, W., and Peters, J. (1996) A hyperthermostable protease of the subtilisin family bound to the surface layer of the archaeon Staphylothermus marinus, *Current biology : CB 6*, 739-749.

[13] Ozbek, S., Engel, J., and Stetefeld, J. (2002) Storage function of cartilage oligomeric matrix protein: the crystal structure of the coiled-coil domain in complex with vitamin D(3), *EMBO J 21*, 5960-5968.

[14] Wong, M. Y., Steudel, Y., and Steudel, R. (2002) Novel species for the sulfur zoo: isomers of S8, *Chemical Physics Letters 364*, 387-392.

[15] Gilson, M. K., Given, J. A., Bush, B. L., McCammon, J. A. (1997) The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review, *Biophys. J. 72*, 1047-1069.

[16] Straatsma, T. P., McCammon, J. A. (1991) Multiconfigurational thermodynamic integration, J. Chem. Phys. 91, 3631-3637.

[17] Boresch, S., Tettinger, F., Leitgeb, M., Karplus, M. (2003) Absolute Binding Free Energies: A Quantitative Approach for their Calculation, *J. Phys. Chem. B 107*, 9535-9551.

[18] Deng, Y., Roux, B. (2006) Calculation of Standard Binding Free Energies: Aromatic Molecules in the T4 Lysozyme L99A Mutant, *J. Chem. Theory Comput. 2*, 1255-1273.

[19] Helms, V., and Wade, R. C. (1998) Hydration Energy Landscape of the Active Site Cavity in Cytochrome P450cam, *Proteins: Structure, Function and Genetics 32*, 381-396

[20] Mobley, D. L., Chodera, J. D., and Dill, K. A. (2006) On the use of orientational restraints and symmetry corrections in alchemical free energy calculations, *J. Chem. Phys. 125*, 084902-1-16

[21] Wang, J., Deng, Y., Roux, B. (2006) Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials, *Biophys. J. 91*, 2798-2814.

[22] Helms, V., and Wade, R. C. (1995) Thermodynamics of Water Mediating Protein-Ligand Interactions in Cytochrome P450cam: A Molecular Dynamics Study, *Biophys. J. 69*, 810-824.

[23] Yin, H., Hummer, G., and Rasaiah, J. C. (2007) Metastable water clusters in the nonpolar cavities of the thermostable protein tetrabrachion, *Journal of the American Chemical Society 129*, 7369-7377.

[24] Ozbek, S., Muller, J. F., Figgemeier, E., and Stetefeld, J. (2005) Favourable mediation of crystal contacts by cocoamidopropylbetaine (CAPB), *Acta crystallographica. Section D, Biological crystallography 61*, 477-480.

[25] Odoh, S. O., Walker, S. M., Meier, M., Stetefeld, J., and Schreckenbach, G. (2011) QM and QM/MM studies of uranyl fluorides in the gas and aqueous phases and in the hydrophobic cavities of tetrabrachion, *Inorganic chemistry 50*, 3141-3152.

[26] Boyd, E. S., and Druschel, G. K. (2013) Involvement of intermediate sulfur species in biological reduction of elemental sulfur under acidic, hydrothermal conditions, *Applied and environmental microbiology 79*, 2061-2068.

[27] Leslie, A. G. W. (1994) *MOSFLM users guide*, MRC-LMB, Cambridge.

[28] CCP4. (1994) Collaborative Computing Project No. 4. (1994). The CCP4 suite: programs for protein crystallography, *Acta Cryst. D50*, 760-763.

[29] Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr 53*, 240-255.

[30] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.

[31] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen V. B., Echols N., Headd, J. J., Hung L. W., Jain, S., Kapral G. J., Grosse Kunstleve, R. W., McCoy A. J., Moriarty N. W., Oeffner, R. D., Read R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures, *Methods 55*, 94-106

[32] Kabsch, W. Xds, Acta Crystallogr D Biol Crystallogr 66, 125-132.

[33] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free, *Journal of computational chemistry 26*, 1701-1718.

[34] Eckert, B., and Steudel, R. (2003) Molecular spectra of sulfur molecules and solid sulfur allotropes, *Top Curr Chem 231*, 31-98.

[35] Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta crystallographica. Section D, Biological crystallography 60*, 1355-1363.

[36] Lee, O. S., Yang, K., Kang, K. D., Koo, I. S., Kim, C. K., and Lee, I. (2004) Ab initio and DFT studies on hydrolyses of phosphorus halides, *Journal of computational chemistry 25*, 1740-1748.

[37] Boulegue, J. (1978) Solubility of Elemental Sulfur in Water at 298K, *Phosphorous and Sulfur and the Related Elements* 5, 127- 128

[38] Meyer, B. (1976) Elemental Sulfur, *Chem. Rev.* 76, 367-388

Chapter 6: Summary and Conclusions

Netrin-4 and Rhodocetin both act by displacing or blocking a partner in a protein complex in an extracellular environment; despite being catalytically inert both proteins are capable of executing dramatic physiological changes. For Netrin-4, the function was not immediately obvious from a homology perspective; structural biology and consequent validation experiments were necessary to elucidate the precise mechanism of action, which was resolved even in the absence of a high resolution structure of the netrin-4/laminin-γ1 complex. Though netrin-4 has a regulatory function on the basement membrane[1] and it modulates of tumour development and angiogenesis[2–4], the precise role it plays is yet unknown. It's expression is irregular among netrins; instead of high expression in brain and nervous tissue, netrin-4 is found primarily in the kidney, eye, olfactory bulb, lung, salivary gland, mammary gland, ovary, and heart[5]. With the possible exception of the ovary and heart, the tissues where netrin-4 is found all have a degree of diffusivity related to their functions, i.e. oxygen and carbon dioxide in the lung and eye, urea and other catabolites in the kidney, etc. Further research will be necessary to determine what part netrin-4 plays in a healthy organism.

In the case of rhodocetin we were able to readily understand its biological function by solving the structure of rhodocetin in complex with its target, integrinα2β1. The structure of the rhodocetin tetramer published in 2009[6] (pdb 3GPR) gives no hint of the large conformational changes necessary for the binding of the α2A domain. Furthermore, the inhibitory complex between the snake venom protein and the integrinα2β1 consists of only the RCγδ dimer and the α2A domain, with the RCαβ dimer released. The resultant inhibitory RCγδ-α2A complex blocks the collagen mediated coagulation pathway by obstructing the access of collagen to its binding site on integrinα2β1. The need for the initial RCαβγδ tetramer is to maintain the solubility of the RCγδ subunit until it has interacted with the α2A domain, as in the absence of RCαβ the integrin binding subunit has a much lower solubility and a tendency to aggregate[7,8]. Thus, a single venom component is able to shut down two major blood clotting pathways and allow the enzymes in the protein cocktail to cause as much tissue damage as possible. Our study implicates a relatively small interaction site on α2-integrin, the loop 2 motif, in the selectivity of RCγδ. Mutation of any of these residues, especially G218 which is absent in α1-integrin, strongly and negatively impacts binding[9].

The S-layer of *S. marinus* is essential for cellular survival. Aside from its role in protecting against thermal and osmotic stress, it’s likely involved in procuring a peptide source for metabolism and a terminal electron acceptor via the unique enzyme, STABLE, and RHCC, respectively. Our results show that a short motif in a much larger protein forms, upon oligomerization, a binding pocket for the uptake of elemental sulfur. The nature of this system may explain why it has been difficult to determine and characterise other sulfur binding systems. Though our study identifies a portion of the S-layer as a potential member of the sulfur sequestering machinery, the details of this process remain undetermined. Examining the protease STABLE may yield more information on the transport of elemental sulfur within the S-layer. The non-covalent complex between STABLE and the tetrabrachion is extremely strong; it persists at 70°C in the presence of 1% SDS[10]. STABLE itself is a poorly characterized protein that consists of 1345 residues and is heavily glycosylated[11]. It has very few identified homologues. About 250 residues are implicated in protease activity, specifically in the peptidase S8 family; the function of the remaining residues remains unclear to date. As the interaction site of STABLE on the tetrabrachion is suspected to be RHCC, a putative role in the sulfur reduction pathway merits further investigation.

Aside from its possible role in sulfur respiration, RHCC is an interesting protein that merits additional study. With its simple fold, high stability, and robust recombinant expression, RHCC has many traits required for a biotech application. While it doesn’t catalyze any industrially useful reaction, nor does the native function of the protein immediately suggest a biomedical application, the small cavities may have advantages in binding ligands of interest to industry over other systems currently in use. For example, due to the hydrophobic nature of the cavities, RHCC has significant potential as a matrix for sequestering low molecular weight polycylic aromatic hydrocarbons (PAHs) from the environment. These compounds are present in crude oil, can escape to the environment during the extraction process or during a spill, and are of significant concern to human health[12]. Industry standard methods of determining the concentration of these compounds relies on either live sampling of biota such as fish or waterfowl and extraction from the liver, or using a passive sampler with a polymer matrix. Both methods have severe drawbacks; live sampling is expensive and imprecise, while the passive sampling devices have long equilibration times (> 14 days) and only weakly bind low molecular weight PAHs[13,14]. However, a passive sample device with a RHCC matrix is capable of much faster equilibration with low MW PAHS and binds in a concentration dependant manner. RHCC may also have potential as a carrier in targeted drug delivery, as demonstrated in a 2015 *Nanomedicine* paper by Klonisch et al., where RHCC carrying cis-platin is capable of selectively killing human glioblastoma cells[15].

While enzymes can be easily categorized by their catalytic function, the functions of non-enzymatic proteins can be much more subtle. In the absence of experimental data, protein functions are often assigned based on sequence or domain conservation, but this can be misleading. In our case there is an excess of literature predicated on the interactions of Net4 with neogenin or other netrin-1 dependence receptors while our data shows that, from a structural perspective, these interactions are unlikely to occur. This assumed kinship between Net4 and Net1 is logical when looking through the lens of homology alone, but structurally conserved proteins can have a multitude of diverse functions and modes of action. Conversely, a simple α-helical homotetramer can yield insights on how archaea can survive and respire under the harshest of conditions. This thesis brings to light new knowledge on three extracellular systems where either the function or mechanism were unclear prior to structural characterization and subsequent validation.

References

1. Reuten, R. *et al.* Structural decoding of netrin-4 reveals a regulatory function towards mature basement membranes. *Nat. Commun.* **7,** 13515 (2016).

2. Lambert, E., Coissieux, M. M., Laudet, V. & Mehlen, P. Netrin-4 acts as a pro-angiogenic factor during zebrafish development. *J. Biol. Chem.* **287,** 3987–3999 (2012).

3. Yuan, Y. *et al.* Netrin-4 is upregulated in breast carcinoma effusions compared to corresponding solid tumors. *Diagn. Cytopathol.* **39,** 562–566 (2011).

4. Nacht, M. *et al.* Netrin-4 regulates angiogenic responses and tumor cell growth. *Exp. Cell Res.* **315,** 784–794 (2009).

5. Li, Y. N. *et al.* The expression and function of netrin-4 in murine ocular tissues. *Exp. Eye Res.* **96,** 24–35 (2012).

6. Eble, J. a *et al.* The alpha2beta1 integrin-specific antagonist rhodocetin is a cruciform, heterotetrameric molecule. *FASEB J.* **23,** 2917–2927 (2009).

7. Navdaev, A., Lochnit, G. & Eble, J. A. The rhodocetin αβ subunit targets GPIb and inhibits von Willebrand factor induced platelet activation. *Toxicon* **57,** 1041–1048 (2011).

8. Eble, J. a & Tuckwell, D. S. The alpha2beta1 integrin inhibitor rhodocetin binds to the A-domain of the integrin alpha2 subunit proximal to the collagen-binding site. *Biochem. J.* **376,** 77–85 (2003).

9. Eble, J. A. *et al.* Dramatic and concerted conformational changes enable rhodocetin to block α2β1 integrin selectively. *PLoS Biol.* **15,** e2001492 (2017).

10. Peters, J. *et al.* Tetrabrachion: A Filamentous Archaebacterial Surface Protein Assembly of Unusual Structure and Extreme Stability. *J. Mol. Biol.* **245,** 385–401 (1995).

11. Mayr, J. *et al.* A hyperthermostable protease of the subtilisin family bound to the surface layer of the archaeon Staphylothermus marinus. *Curr. Biol.* **6,** 739–49 (1996).

12. Poursafa, P. *et al.* A Systematic Review on the Effects of Polycyclic Aromatic Hydrocarbons on Cardiometabolic Impairment. *Int. J. Prev. Med.* **8,** 19 (2017).

13. Allan, I. J., Harman, C., Kringstad, A. & Bratsberg, E. Effect of sampler material on the uptake of PAHs into passive sampling devices. *Chemosphere* **79,** 470–475 (2010).

14. Saija, E. *et al.* Determination and quantification of PCBs, POCs and PAHs in Thunnus thynnus from the Straits of Messina (Italy). *Data Br.* **7,** 129–134 (2016).

15. Thanasupawat, T. *et al.* Platinum (IV) coiled coil nanotubes selectively kill human glioblastoma cells. *Nanomedicine Nanotechnology, Biol. Med.* **11,** 913–925 (2015).

Appendix I: Additional Supporting Information for Structural decoding of netrin-4 reveals unique non-enzymatic disruptive forces towards mature basement membranes.

Statement of Study Approval

All animal experiments were performed in accordance with the German laws for animal protection and were approved by the State Office North Rhine-Westphalia (LANUV), Germany.

Tables

Table I-1. Hydrodynamic data for Net4-ΔC, LNγ1, Net4-ΔC-γ1LN-LEa1-4, and Net4-FL-γ1LN-LEa1-4

Parameters	Net4-ΔC	Net4-FL	γ1LN-LEa1-4		Net4-ΔC-γ1LN-LEa1-4		Net4-FL-γ1LN-LEa1-4	
	Exp.[d]	Exp.	Exp.	CORAL[e]	Exp.	SASREF[e]	Exp.	SASREF[e]
r_H (nm)[a]	4.60±0.20	5.76 ± 0.02	4.02 ± 0.05	3.95 ± 0.08	5.40 ± 0.10	5.37 ± 0.02	5.77 ± 0.02	5.86 ± 0.02
r_G (nm)[b]	4.42±0.20	6.03±0.06	4.20 ± 0.04	4.12 ± 0.01	5.31 ± 0.07	5.31 ± 0.05	6.07 ± 0.04	6.00 ± 0.04
D_{max} (nm)[b]	16.00	20.0	14.1	14.2 ± 0.27	17.8	20.4 ± 0.28	19.0	21.5 ± 0.25
$S^0_{20,w}$ (S)[c]	3.51 ± 0.07	4.23 ± 0.06[f] 6.82 ± 0.41[g]	3.62 ± 0.04	3.71 ± 0.10	5.11 ± 0.04	5.32 ± 0.02	5.15 ± 0.01	5.55 ± 0.10
χ[h]	-	-	-	1.08	-	1.14	-	1.45
NSD[i]	-	-	-	0.91 ± 0.12	-	0.27 ± 0.03	-	0.35 ± 0.05

[a] Experimentally determined from DLS data, with error obtained from linear regression analysis

[b] Experimentally determined from $p(r)$ analysis by GNOM

[c] Experimentally determined using Analytical Ultracentrifuge

[d] Previously published data[1]

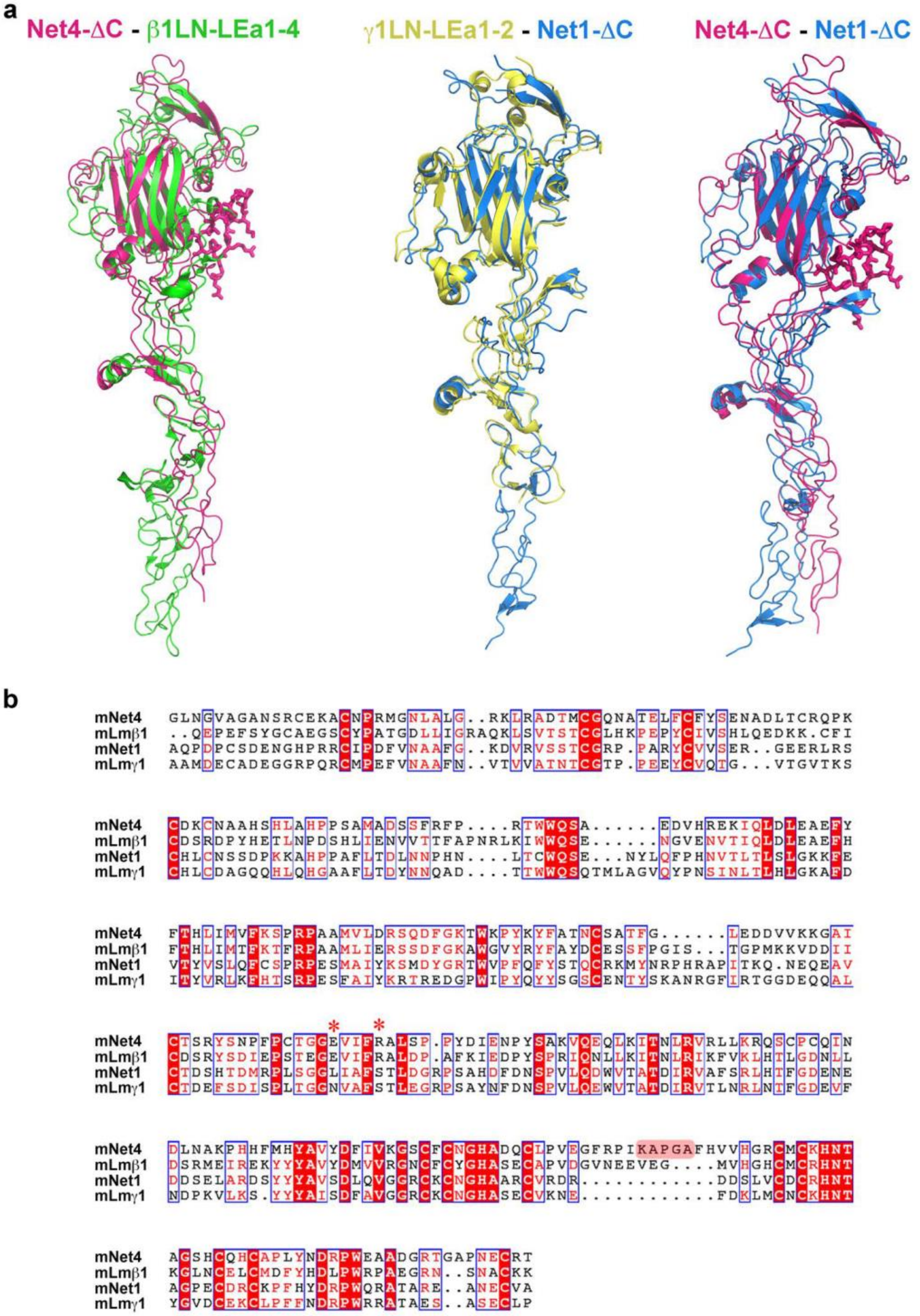
[e] Model-based parameters calculated from HYDROPRO, with errors obtained as the standard deviation from multiple models.

[f, g] Peak 1 and 2 respectively from Sedimentation coefficient distribution

[h] Goodness of fit parameter that represents agreement between SAXS raw data and data back calculated from models

[i] Normalized spatial discrepancy parameter representing agreement between multiple models

Figures



```
mNet4   GLNGVAGANSRCEKACNPRMGNLALG..RKLRADTMCGQNATELFCFYSENADLTCRQPK
mLmβ1   ..QEPEFSYGCAEGSCYPATGDLLIGRAQKLSVTSTCGLHKPEPYCIVSHLQEDKK.CFI
mNet1   AQPDPCSDENGHPRRCIPDFVNAAFG..KDVRVSSTCGRP.PARYCVVSER..GEERLRS
mLmγ1   AAMDECADEGGRPQRCMPEFVNAAFN..VTVVATNTCGTP.PEEYCVQTG...VTGVTKS

mNet4   CDKCNAAHSHLAHPPSAMADSSFRFP....RTWWQSA......EDVHREKIQLDLEAEFY
mLmβ1   CDSRDPYHETLNPDSHLIENVVTTFAPNRLKIWWQSE......NGVENVTIQLDLEAEFH
mNet1   CHLCNSSDPKKAHPPAFLTDLNNPHN....LTCWQSE...NYLQFPHNVTLTLSLGKKFE
mLmγ1   CHLCDAGQQHLQHGAAFLTDYNNQAD....TTWWQSQTMLAGVQYPNSINLTLHLGKAFD

mNet4   FTHLIMVFKSPRPAAMVLDRSQDFGKTWKPYKYFATNCSATFG......LEDDVVKKGAI
mLmβ1   FTHLIMTFKTFRPAAMLIERSSDFGKAWGVYRYFAYDCESSFPGIS...TGPMKKVDDII
mNet1   VTYVSLQFCSPRPESMAIYKSMDYGRTWVPFQFYSTQCRKMYNRPHRAPITKQ.NEQEAV
mLmγ1   ITYVRLKFHTSRPESFAIYKRTREDGPWIPYQYYSGSCENTYSKANRGFIRTGGDEQQAL

                      *   *
mNet4   CTSRYSNPFPCTGGEVIFRALSP.PYDIENPYSAKVQEQLKITNLRVRLLKRQSCPCQIN
mLmβ1   CDSRYSDIEPSTEGEVIFRALDP.AFKIEDPYSPRIQNLLKITNLRIKFVKLHTLGDNLL
mNet1   CTDSHTDMRPLSGGLIAFSTLDGRPSAHDFDNSPVLQDWVTATDIRVAFSRLHTFGDENE
mLmγ1   CTDEFSDISPLTGGNVAFSTLEGRPSAYNFDNSPVLQEWVTATDIRVTLNRLNTFGDEVF

mNet4   DLNAKPHHFMHYAVYDFIVKGSCFCNGHADQCLPVEGFRPIKAPGAFHVVHGRCMCKHNT
mLmβ1   DSRMEIREKYYYAVYDMVVRGNCFCYGHASECAPVDGVNEEVEG....MVHGHCMCRHNT
mNet1   DDSELARDSYYYAVSDLQVGGRCKCNGHAARCVRDR............DDSLVCDCRHNT
mLmγ1   NDPKVLKS.YYYAISDFAVGGRCKCNGHASECVKNE............FDKLMCNCKHNT

mNet4   AGSHCQHCAPLYNDRPWEAADGRTGAPNECRT
mLmβ1   KGLNCELCMDFYHDLPWRPAEGRN..SNACKK
mNet1   AGPECDRCKPFHYDRPWQRATARE..ANECVA
mLmγ1   YGVDCEKCLPFFNDRPWRRATAES..ASECLP
```

Figure I-1. Alignment of netrins and laminins.

(**a**) 3D structure comparison of Net4-ΔC-β1LN-LEa1-4, γ1LN-LEa1-2-Net1-ΔC, and Net4-ΔC-Net1ΔC. (**b**) The LN and LE1 domain without the signal peptide of mouse Net4 (mNet4, NP_067295), mouse laminin β1 (mLmβ1, NP_032508), mouse Net1 (mNet1, NP_032770), and mouse laminin γ1 (mLmγ1, NP_034813) were aligned using T-coffee[2] and formatted with ESPRIPT. Asterisks indicate the amino acids involved in binding between Net4 and Lmγ1. Interestingly, the Net4 E195 and R199 are conserved in Lmβ1, thought the KAPGA loop is unique to Net4.

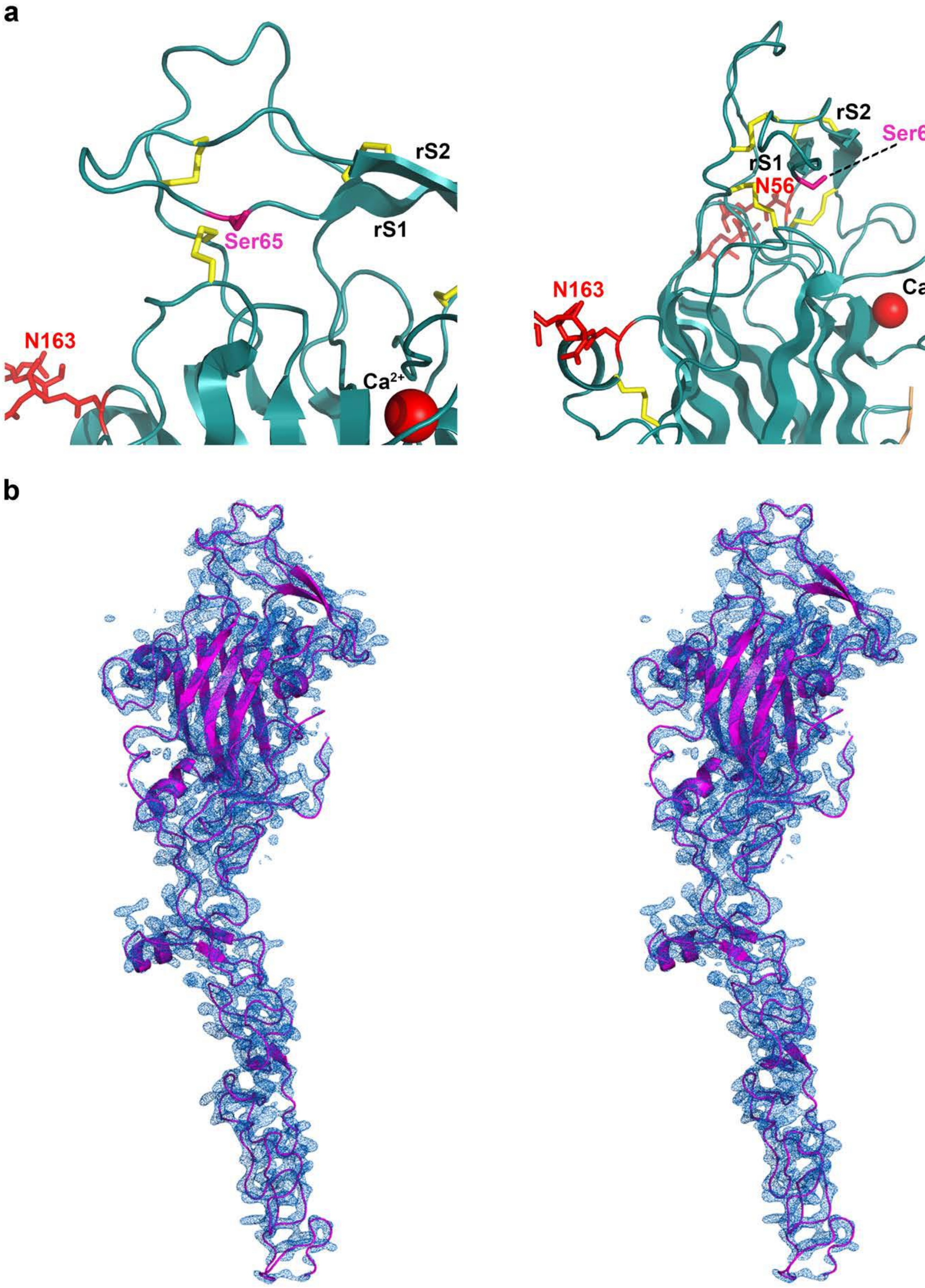


Figure I-2. Structural features of Net4-ΔC.

(**a**) The conserved Ser65 is held in place by 2 unique disulfide bridges. The loop between rS1 and rS2 in the LN domain of Net4 is held in place by disulfide bridges 72-236 and 191-234, which are absent in all netrin and laminin structures published to date. In Lmβ1, the replacement of this conserved serine by arginine allows the formation of a binary complex between β1 and γl laminins; however, the ternary complex formation with the αl laminin is abolished. (**b**) Stereo image of electron density of Net4-ΔC. Net4-ΔC is represented in cartoon mode with secondary structure elements highlighted and sidechains omitted for simplicity. The 2FoFc map is contoured at 2.0σ.

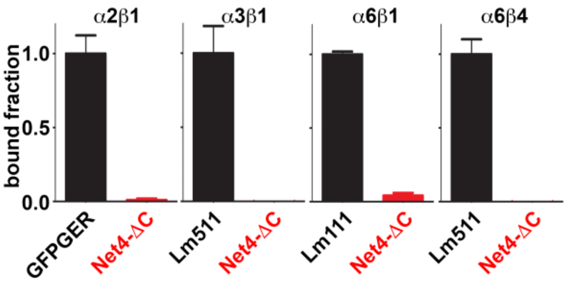
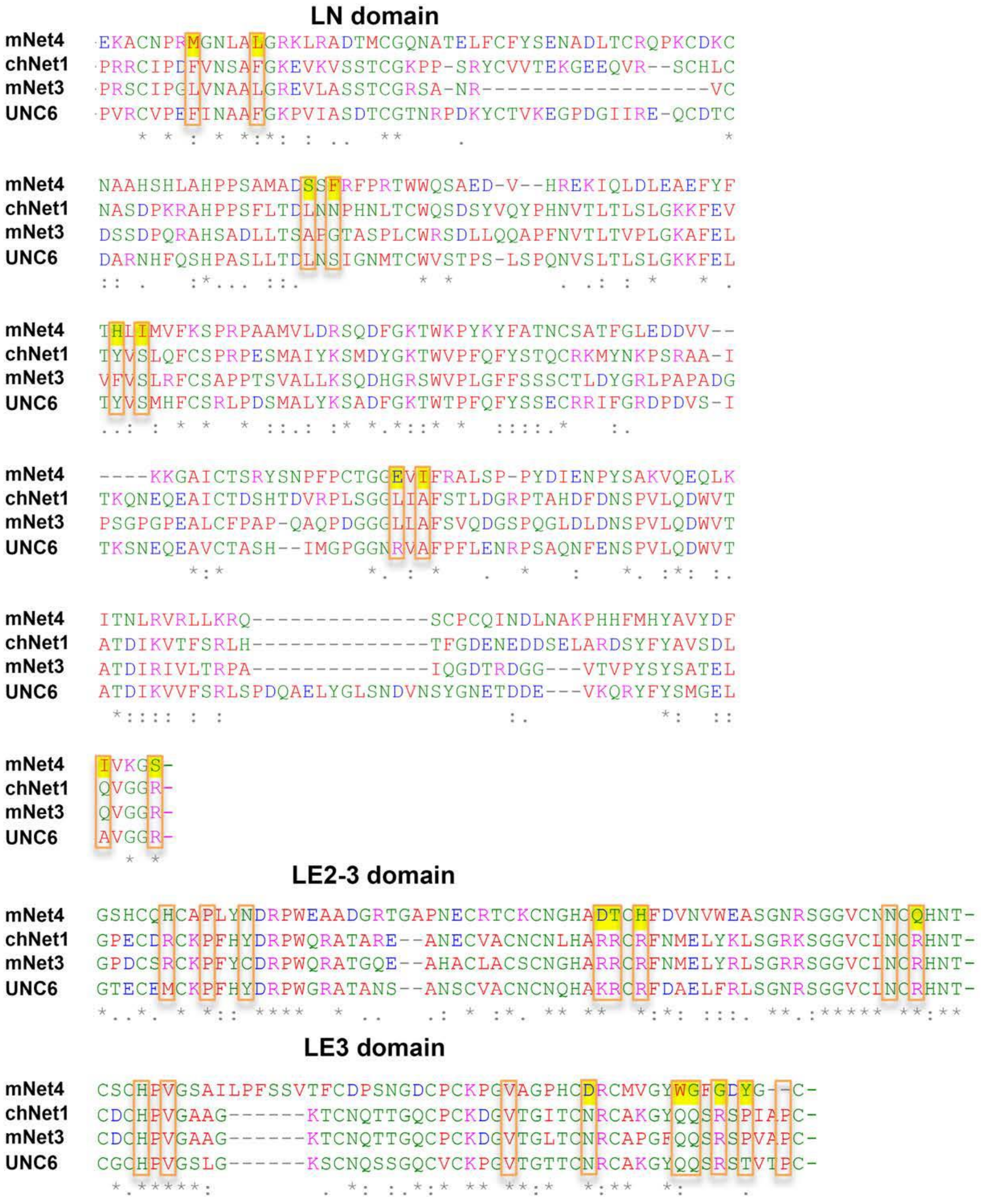


Figure I-3. Net4 binding to integrins.

Solid-phase binding studies of Net4-ΔC (analyte) binding to immobilized integrin α2β1, α3β1, α6β1, and α6β4. The GFPGER peptide, laminin 511 (Lm511), and laminin 111 (LM111) were used as positive controls. Error bars, s.d. (n = 3 independent technical replicates.



```
                    LN domain
mNet4   EKACNPRMGNLALGRKLRADTMCGQNATELFCFYSENADLTCRQPKCDKC
chNet1  PRRCIPDFVNSAFGKEVKVSSTCGKPP-SRYCVVTEKGEEQVR--SCHLC
mNet3   PRSCIPGLVNAALGREVLASSTCGRSA-NR------------------VC
UNC6    PVRCVPEFINAAFGKPVIASDTCGTNRPDKYCTVKEGPDGIIRE-QCDTC
           * * : * *:*: : ..  **    .                    *

mNet4   NAAHSHLAHPPSAMADSSFRFPRTWWQSAED-V--HREKIQLDLEAEFYF
chNet1  NASDPKRAHPPSFLTDLNNPHNLTCWQSDSYVQYPHNVTLTLSLGKKFEV
mNet3   DSSDPQRAHSADLLTSAFGTASPLCWRSDLLQQAPFNVTLTVPLGKAFEL
UNC6    DARNHFQSHPASLLTDLNSIGNMTCWVSTPS-LSPQNVSLTLSLGKKFEL
        :: .   :*... ::.         * *        . .: : *   * .

mNet4   THLIMVFKSPRPAAMVLDRSQDFGKTWKPYKYFATNCSATFGLEDDVV--
chNet1  TYVSLQFCSPRPESMAIYKSMDYGKTWVPFQFYSTQCRKMYNKPSRAA-I
mNet3   VFVSLRFCSAPPTSVALLKSQDHGRSWVPLGFFSSSCTLDYGRLPAPADG
UNC6    TYVSMHFCSRLPDSMALYKSADFGKTWTPFQFYSSECRRIFGRDPDVS-I
        ..: : * *  * ::.: :* *.*::* *  ::::.*   :.

mNet4   ----KKGAICTSRYSNPFPCTGGEVIFRALSP-PYDIENPYSAKVQEQLK
chNet1  TKQNEQEAICTDSHTDVRPLSGGLIAFSTLDGRPTAHDFDNSPVLQDWVT
mNet3   PSGPGPEALCFPAP-QAQPDGGGLLAFSVQDGSPQGLDLDNSPVLQDWVT
UNC6    TKSNEQEAVCTASH--IMGPGGNRVAFPFLENRPSAQNFENSPVLQDWVT
               *:*           *. : *   .  *   :   *. :*: :.

mNet4   ITNLRVRLLKRQ--------------SCPCQINDLNAKPHHFMHYAVYDF
chNet1  ATDIKVTFSRLH--------------TFGDENEDDSELARDSYFYAVSDL
mNet3   ATDIRIVLTRPA--------------IQGDTRDGG---VTVPYSYSATEL
UNC6    ATDIKVVFSRLSPDQAELYGLSNDVNSYGNETDDE---VKQRYFYSMGEL
         *:::: : :                      :.          *:  ::

mNet4   IVKGS-
chNet1  QVGGR-
mNet3   QVGGR-
UNC6    AVGGR-
         * *

                    LE2-3 domain
mNet4   GSHCQHCAPLYNDRPWEAADGRTGAPNECRTCKCNGHADTCHFDVNVWEASGNRSGGVCNNCQHNT-
chNet1  GPECDRCKPFHYDRPWQRATARE--ANECVACNCNLHARRCRFNMELYKLSGRKSGGVCINCRHNT-
mNet3   GPDCSRCKPFYCDRPWQRATGQE--AHACLACSCNGHARRCRFNMELYRLSGRRSGGVCINCRHNT-
UNC6    GTECEMCKPFHYDRPWGRATANS--ANSCVACNCNQHAKRCRFDAELFRLSGNRSGGVCINCRHNT-
        *..*. * *:: ****  * ..   .: * :*. ** **  *:*: :::. **.:***** **:**

                    LE3 domain
mNet4   CSCHPVGSAILPFSSVTFCDPSNGDCPCKPGVAGPHCDRCMVGYWGFGDYG--C-
chNet1  CDCHPVGAAG------KTCNQTTGQCPCKDGVTGITCNRCAKGYQQSRSPIAPC-
mNet3   CDCHPVGAAG------KTCNQTTGQCPCKDGVTGLTCNRCAPGFQQSRSPVAPC-
UNC6    CGCHPVGSLG------KSCNQSSGQCVCKPGVTGTTCNRCAKGYQQSRSTVTPC-
        *.*****:        . *: :.*:* ** **:*  *:**  *:    .
```

Figure I-4. Alignment of netrins and laminins.

Sequence comparison between Net4 (mNet4, NP_067295), Net1 (chNetl, NP_990750), netrin-3 (mNet3, NP_035077), and unc-6 (UNC6, NP_509165). Epitopes which are involved in DCC/neogenin and UNC5 binding are encircled. The non-conserved amino acids within the mNet4 sequence are highlighted in yellow.

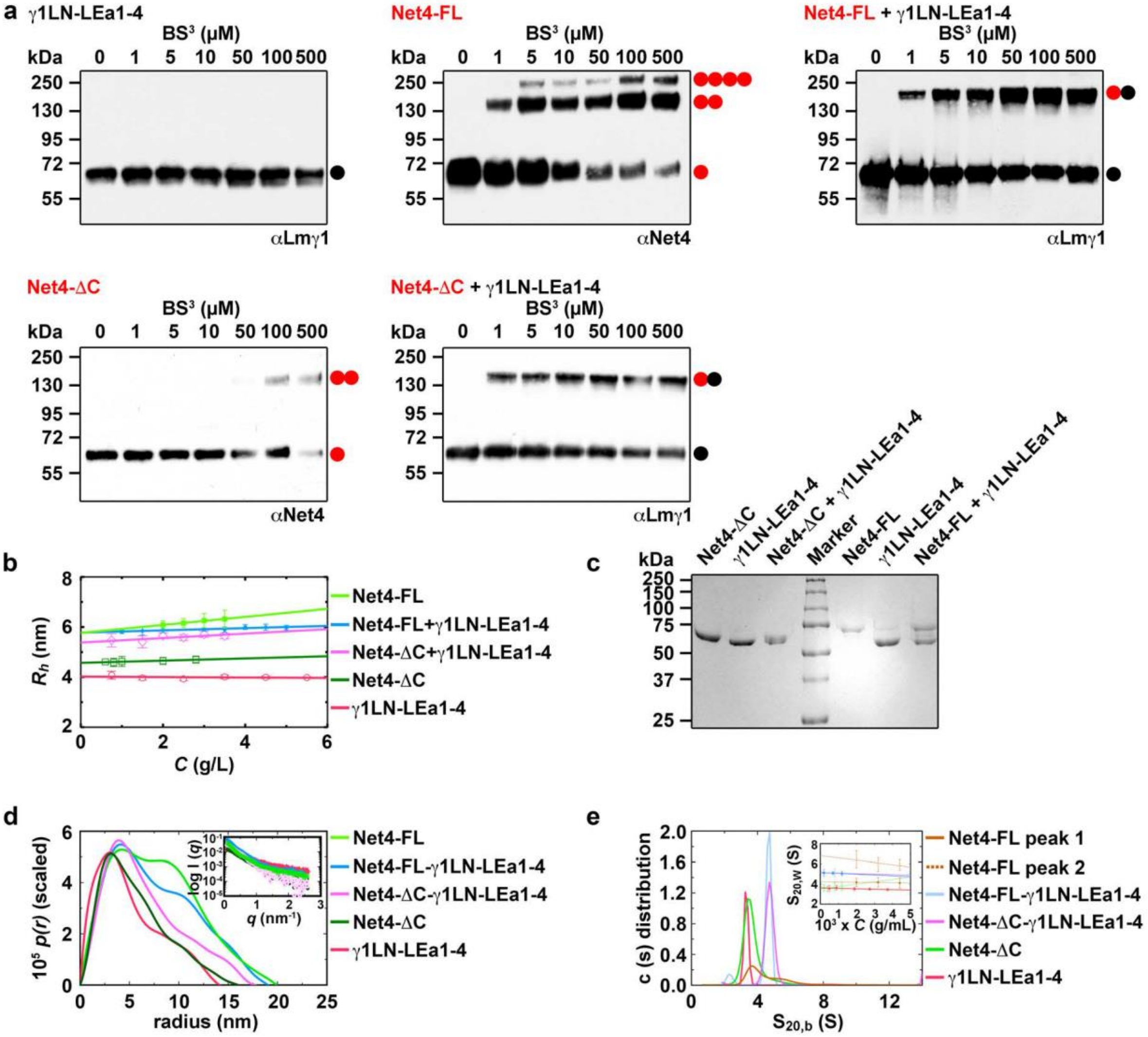


Figure I-5. Laminin-γ1 interferes with the Net4-FL dimer resulting in a 1:1 complex.

(**a**) Cross-linking studies of γ1LN-LEa1-4, Net4-FL, and Net4-FL in complex with γ1LN-LEa1-4. Cross-linking studies were performed using BS3 with increasing concentrations from 0 to 500 pM and detected by either specific Net4 or Lmγ1 antibodies. The western blot represents the presence of monomeric γ1LN-LEa1-4 at all cross-linker concentrations (top left). Net4-FL reveals dimers and tetramers upon increase in BS3 concentration (top middle). Western blot suggests that upon addition of γ1LN-LEa1-4, Net4-FL forms a 1:1 complex with γ1LN-LEa1-4 over the whole concentration regime (top right). Data for Net4-ΔC suggests the presence of dimers at high cross-linker concentrations, whereas Net4-ΔC together with γ1LN-LEa1-4 reveals a stable heterodimeric complex. (**b**) Dynamic light scattering analysis for Net4-ΔC, Net4-FL, γ1LN-LEa1-4 and their complexes. Data for each species were collected at concentrations ranging from 0.75 to 5 mg mL^{-1} and the hydrodynamic radius from each concentration (y-axis) was plotted against concentration (x-axis). Error bars, s.d. for each concentration (n = 3 independent technical replicates). (**c**) SDS-PAGE analysis for SEC purified Net4-ΔC (lane 1), γ1 LN-LEa1-4 (lane 2 and 6), Net4-FL (lane 5) and a complex of Net4-ΔC-γ1LN-LEa1-4 purified from SEC (lane 3) as well as the complex of Net4-FL-γ1LN-LEa1-4 (lane 7) revealed the formation of stable complexes. (**d**) The p(r) distribution function analysis for Net4-ΔC, Net4-FL, γ1 LN-LEa1-4 and a complex of Net4-ΔC-γ1 LN-LEa1-4 as well as Net4-FL-γ1LN-LEa1-4 with raw data (inset) indicating maximum particle dimension on X-axis. (**e**) Sedimentation coefficient distribution at single concentration for Net4-ΔC (1.2 mg mL^{-1}), γ1LN-LEa1-4 (1.9 mg mL^{-1}), Net4-FL (2.0 mg mL^{-1}), Net4-ΔC-γ1LN-LEa1-4 (1.2 mg mL^{-1}) and Net4-FL-γ1LN-LEa1-4 (1.2 mg mL^{-1}) complex indicating that, with the exception of Net4-FL, all species are highly monodisperse. In case of Net4-FL two species are clearly visible corresponding to monomer and dimer peak. The concentration dependence of the sedimentation coefficient for each species with standard deviation is present Table 6-1.

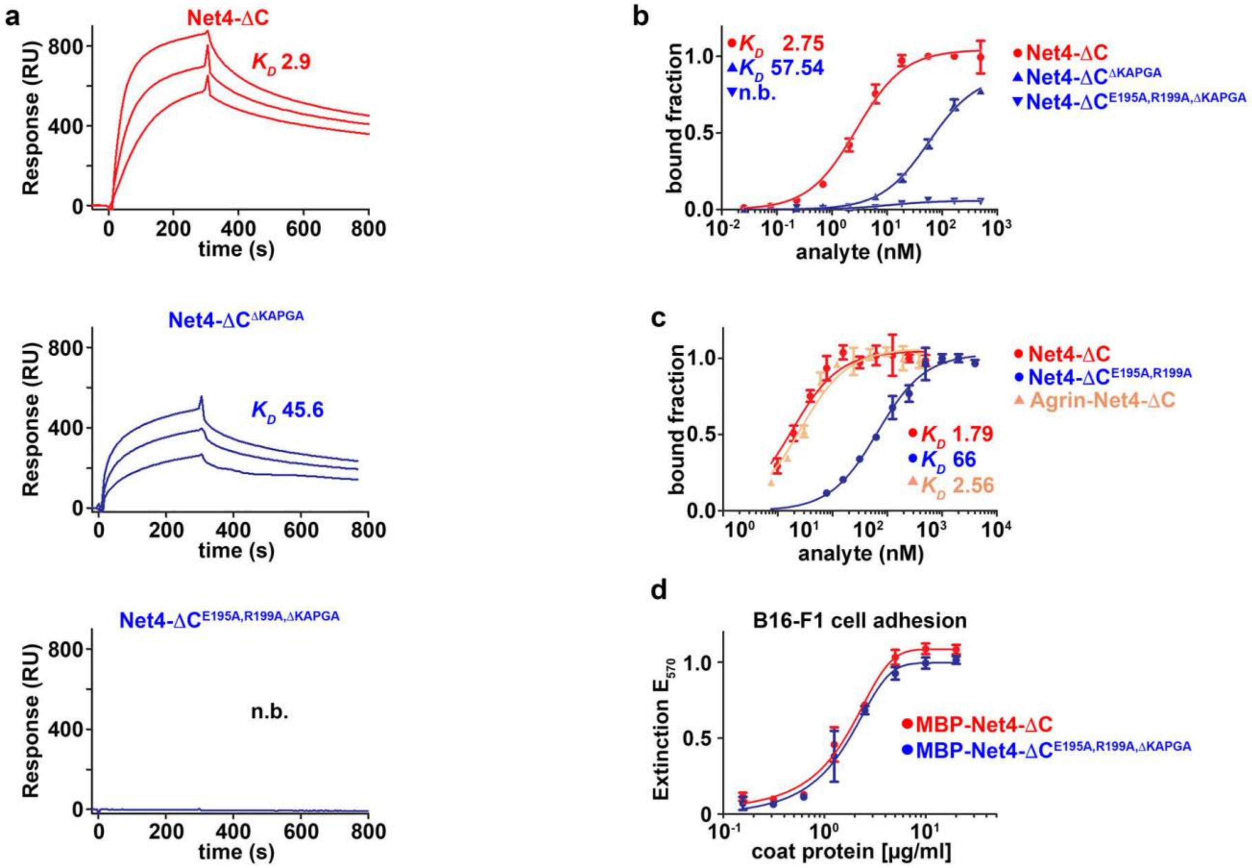


Figure I-6. Identification of the Lmγ1 binding epitopes in Net4.

(**a**) SPR binding studies of Net4-ΔC, Net4-ΔC$^{\Delta KAPGA}$, and Net4-ΔC$^{E195A,R199A,\Delta KAPGA}$ (concentrations: 62.5, 31.25, and 15.625 nM) to immobilized γ1LN-LEa1-4. (**b**) Solid-phase binding studies of Net4-ΔC, Net4-ΔC$^{\Delta KAPGA}$, and Net4-ΔC$^{E195A,R199A,\Delta KAPGA}$ (analyte) binding to immobilized γ1LN-LEa1-4. (**c**) Solid-phase binding studies of Net4-ΔC, Net4-ΔC^{E195A,R199A}, and Agrin-Net4-ΔC (analyte) binding to immobilized γ1LN-LEa1-4. (**d**) Cell adhesion assay of mouse melanoma B16-F1 cells to immobilized MBP-Net4-ΔC and MBP-Net4-ΔC$^{E195A,R199A,\Delta KAPGA}$. K_D values (a-c) are shown in the graph (n.b., no binding). Error bars, s.d. (n = 3 independent technical replicates).

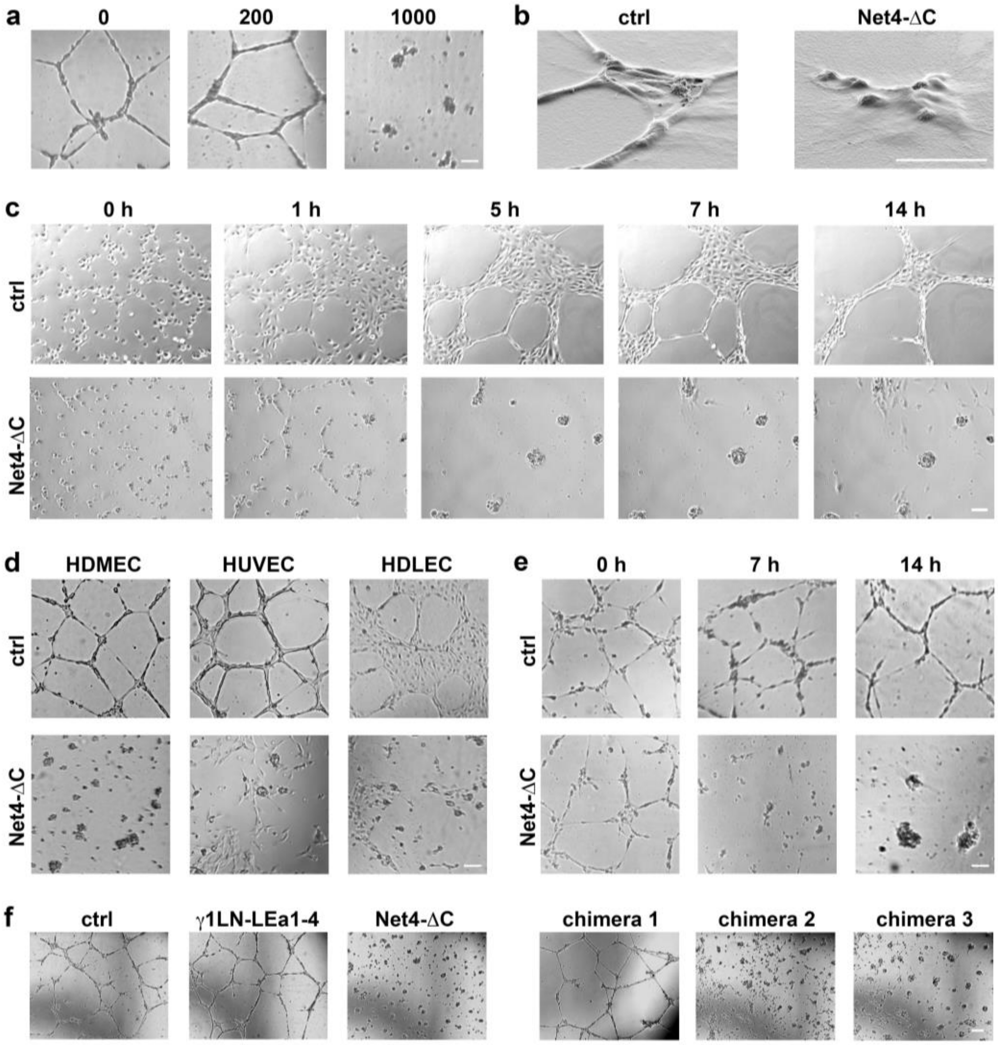


Figure I-7. The laminin network within Matrigel is necessary to establish and maintain endothelial tube-like structures.

(**a**) Treatment of HDMEC cells forming tube-like structures on Matrigel with different concentrations of Net4-ΔC (0-1000 nM). (**b**) Scanning electron microscopy images of untreated and Net4-ΔC (1 pM) treated HDMEC cells on Matrigel. (**c**) Tube formation images of untreated and Net4-ΔC -treated (1 pM) HDMEC cells at different time points (0, 1, 5. 7. and 14 h). (**d**)

Human dermal microvascular endothelial cells (HDMECs), human umbilical vein endothelial cells (HUVECs), and human dermal lymphatic endothelial cells (HDLECs) were treated with Net4-ΔC (1 pM) on Matrigel. (**e**) Preformed HDMEC tube-like structures (7 h after seeding) were treated with and without Net4-ΔC (1 pM) and images were taken from 0, 7, and 14 h post treatment. (**f**) Determination of the Net4 domains mediating the inhibition of HDMEC formed tube-like structures on Matrigel. Net4-ΔC completely inhibits the formation of tubes as well as Net4 domain swap mutants containing the laminin γ1 binding region LN-LE1 indicated by the inhibiting activity of the swap mutants Net4LN-LE2 - γ1LEa3-4 and Net4LN-LE1 - γ1 LEa2-4. Treatment of endothelial cells with the laminin fragment γ1LN-LEa1-4 or the swap mutant γ1LN-LEa1 - Net4LE2-31/2 revealed no difference compared to untreated HDMECs. Scale bars, (a and c-f) 100 pm, (b) 50 pm.

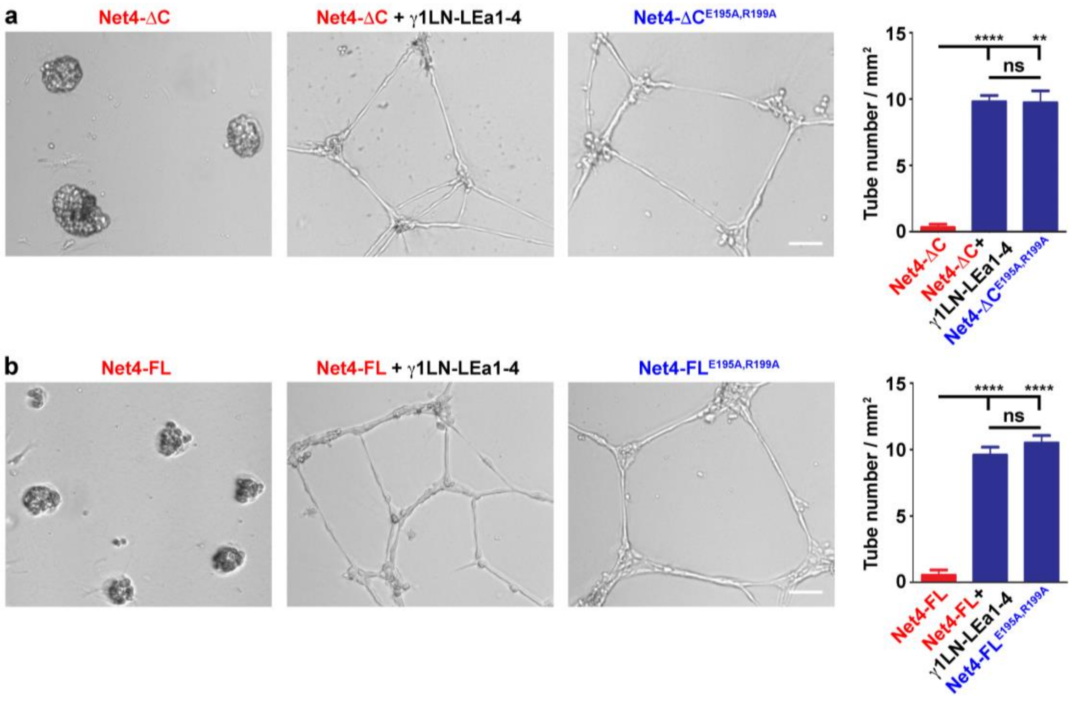


Figure I-8. Inhibition of endothelial tube-like structures through Net4.

The concentration of all used proteins is 1 pM. (**a**) HDMECs were treated with Net4-ΔC, Net4-$\Delta C^{E195A,R199A}$ in combination with γ1LN-LEa1-4 and Net4-ΔC. Statistical analysis of the tube number per mm^2 (mean ± s.d.; n = 3; **P = 0.0015, ****P = 0.00006. *P*- values from two sided *t*-test). (**b**) HDMECs were treated with Net4-FL, Net4-FL in combination with γ1LN-LEa1-4, and Net4-FL^{E195A,R199A}. Statistical analysis of the tube number per mm^2 (mean ± s.d.; n = 3; ****P = 0.00008 (Net4-FL vs. Net4-FL + γ1LN-LEa1-4), ****P = 0.00004 (Net4-FL vs. Net4-FL^{E195A,R199A})). Scale bar. (a and b) 100 pm. Error bars, s.d. (n = 3 independent cell cultures).

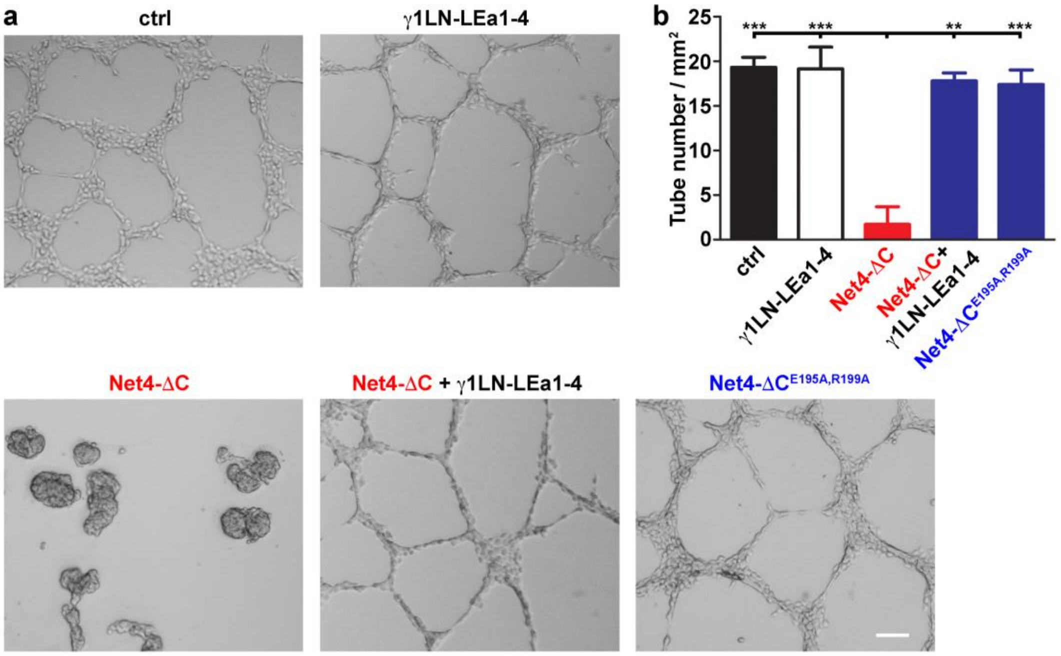


Figure I-9. Inhibition of B16-F1 tube-like structure formation through Net4.

(**a**) B16-F1 cells seeded on Matrigel were treated with 1 pM γ1LN-LEa1-4, Net4-ΔC, Net4-ΔC in combination with γ1LN-LEa1-4, and the laminin-binding mutant Net4-ΔC^{E195AR199A}. (**b**) Statistical analysis of tube number per mm^2 (mean ± s.d.; n = 3; ***P = 0.0008 (ctrl vs. Net4-ΔC), ***P = 0.0008 (γ1LN-LEa1-4 vs. Net4-ΔC), **P = 0.0014 (γ1LN-LEa1-4 + Net4-ΔC vs. Net4-ΔC), ***P = 0.0005 (Net4-ΔC^{E195A,R199A} vs. Net4-ΔC)). Scale bar, (a) 100 pm. Error bars, s.d. (n = 3 independent cell cultures).

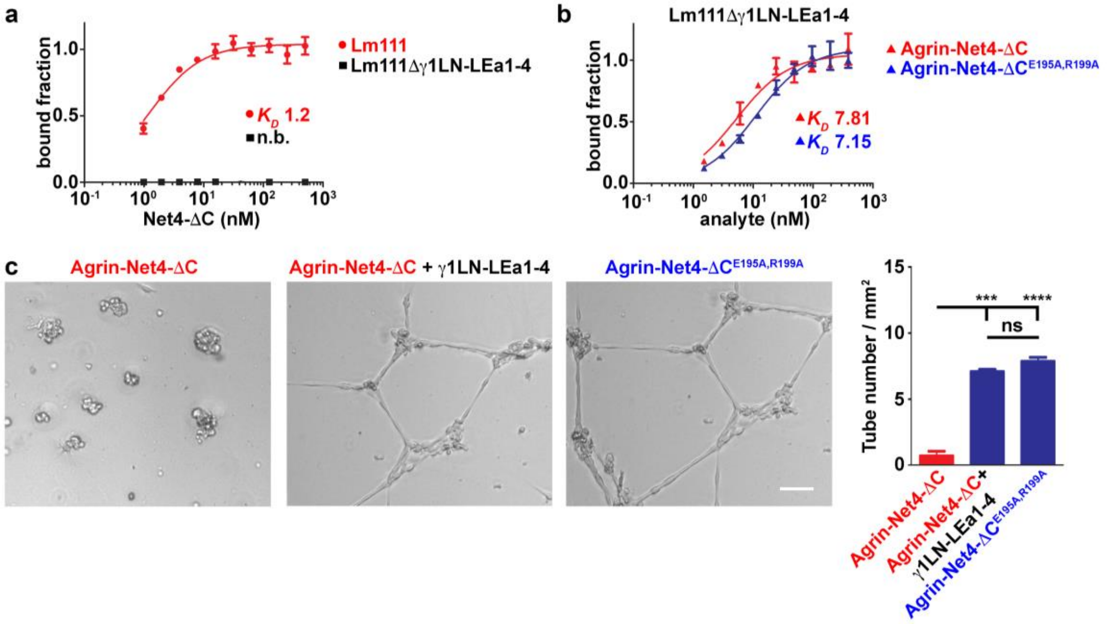
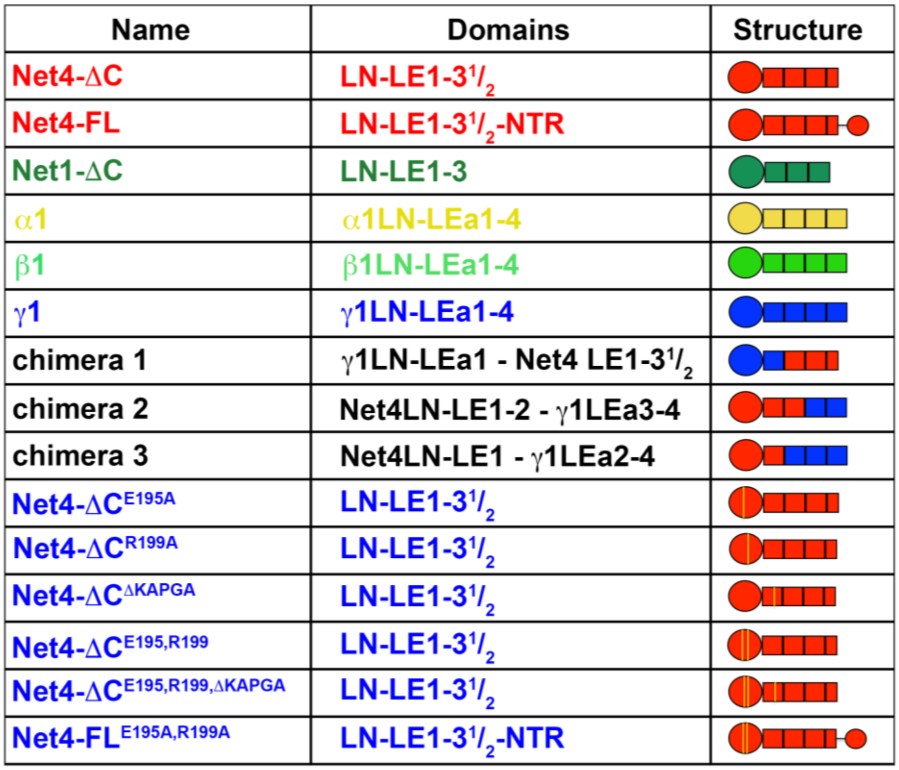


Figure I-10. Inhibitory activity of Net4 is independent from matrix deposition.

(**a**) Solid-phase binding studies of Net4-ΔC, Net4-ΔC^{E195A,R199A}, and Agrin-Net4-ΔC (analyte) binding to immobilized γ1LN-LEa1-4. (**b**) Solid-phase binding studies of Net4-ΔC (analyte) binding to immobilized laminin 111 (Lm111) and Lm111 with a truncation of the N-terminal γ1LN-LEa1-4 domains (Lm111Δγ1LN-LEa1-4). K_D values are shown in the graph (n.b., no binding). Error bars (a and b), s.d. (n = 3 independent technical replicates). (**c**) HDMECs were treated with 1 pM Agrin-Net4-ΔC, Agrin-Net4-ΔC in combination with γ1LN-LEa1-4, and Agrin-Net4-ΔC^{E195A,R199A}. Statistical analysis of the tube number per mm^2 (mean ± s.d.; n = 3; ***P = 0.00013, ****P = 0.000013). Scale bar, 100 pm. Error bars, s.d. (n = 3 independent cell cultures).



Name	Domains	Structure
Net4-ΔC	LN-LE1-$3^1/_2$	
Net4-FL	LN-LE1-$3^1/_2$-NTR	
Net1-ΔC	LN-LE1-3	
α1	α1LN-LEa1-4	
β1	β1LN-LEa1-4	
γ1	γ1LN-LEa1-4	
chimera 1	γ1LN-LEa1 - Net4 LE1-$3^1/_2$	
chimera 2	Net4LN-LE1-2 - γ1LEa3-4	
chimera 3	Net4LN-LE1 - γ1LEa2-4	
Net4-ΔC^{E195A}	LN-LE1-$3^1/_2$	
Net4-ΔC^{R199A}	LN-LE1-$3^1/_2$	
Net4-ΔC$^{\Delta KAPGA}$	LN-LE1-$3^1/_2$	
Net4-ΔC^{E195,R199}	LN-LE1-$3^1/_2$	
Net4-ΔC$^{E195,R199,\Delta KAPGA}$	LN-LE1-$3^1/_2$	
Net4-FL^{E195A,R199A}	LN-LE1-$3^1/_2$-NTR	

Figure I-11. Overview of recombinant proteins used in this study.

References

1. Patel, T.R. et al. Determination of a molecular shape for netrin-4 and small angle X-ray scattering measurements. *Matrix biol* **31,** 135-40 (2012)

2. Notredame, C. et al. T Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302,** 205-17 (2000)

Appendix II: Additional Supporting Information for Dramatic and concerted conformational changes enable rhodocetin to block α2β1 integrin selectively

Acknowledgements

This work was financially supported by the Deutsche Forschungsgemeinschaft (DFG) via the SFB1009 project A09 (JAE) and project Eb177/13-1 (JAE), and by the Heart and Stroke Foundation of Canada (G-14-000625) (JS), NSERC-Research Tool and Infrastructure (JS), and the Canada Research Chair Program (JS). We thank Barbara Schedding, Margret Bahl, Marion Berthold and Alletta Schmidt-Hederich for their technical assistance. We thank Dr. T. Bracht for having isolated several monoclonal antibodies against rhodocetin. We also appreciate the generous gift of JA202 antibody from Dr. D. Tuckwell. We would also like to thank the support staff at the Canadian Light Source CMCF 08.ID1 beamline for assistance with data collection.

Tables

Table II-1. PCR primers used to generate integrin mutants

Outer primers:	
Forward outer primer: (*NdeI* site underlined)	5’-GCAGCCATATGGGAGGTTCTCCTTCCCTCATAGATGTTGTGGTTGTG-3’
Reverse outer primer: (*BamHI* site underlined)	5’-AGCCGGATCCTCGAGCTACTAACCTTCAATGCTGAAAATTTGTTC-3’
Inner primers: (mutation sites are underlined)	
S214A-forward:	5’-GCAACATCCCAGACAGCTCAATATGGTGGGG-3’
S214A-reverse:	5’-CCCCACCATATTGAGCTGTCTGGGATGTTGC-3’

Y216G-forward:	5’-CCCAGACATCCCAAGGTGGTGGGGACCTCAC-3’
Y216G-reverse:	5’-GTGAGGTCCCCACCACCTTGGGATGTCTGGG-3’
G217K-forward:	5’-CAGACATCCCAATATAAAGGGGACCTCACAAAC-3‘
G217K-reverse:	5’-GTTTGTGAGGTCCCCTTTATATTGGGATGTCTG-3‘
G218L-forward:	5’-GACATCCCAATATGGTCTGGACCTCACAAACAC-‘
G218L-reverse	5’-GTGTTTGTGAGGTCCAGACCATATTGGGATGTC-3’
D219A-forward:	5’-CAATATGGTGGGGCACTCACAAACACATTCGGAGC-3‘
D219A-reverse:	5’-GCTCCGAATGTGTTTGTGAGTGCCCCACCATATTG-3’

Figures

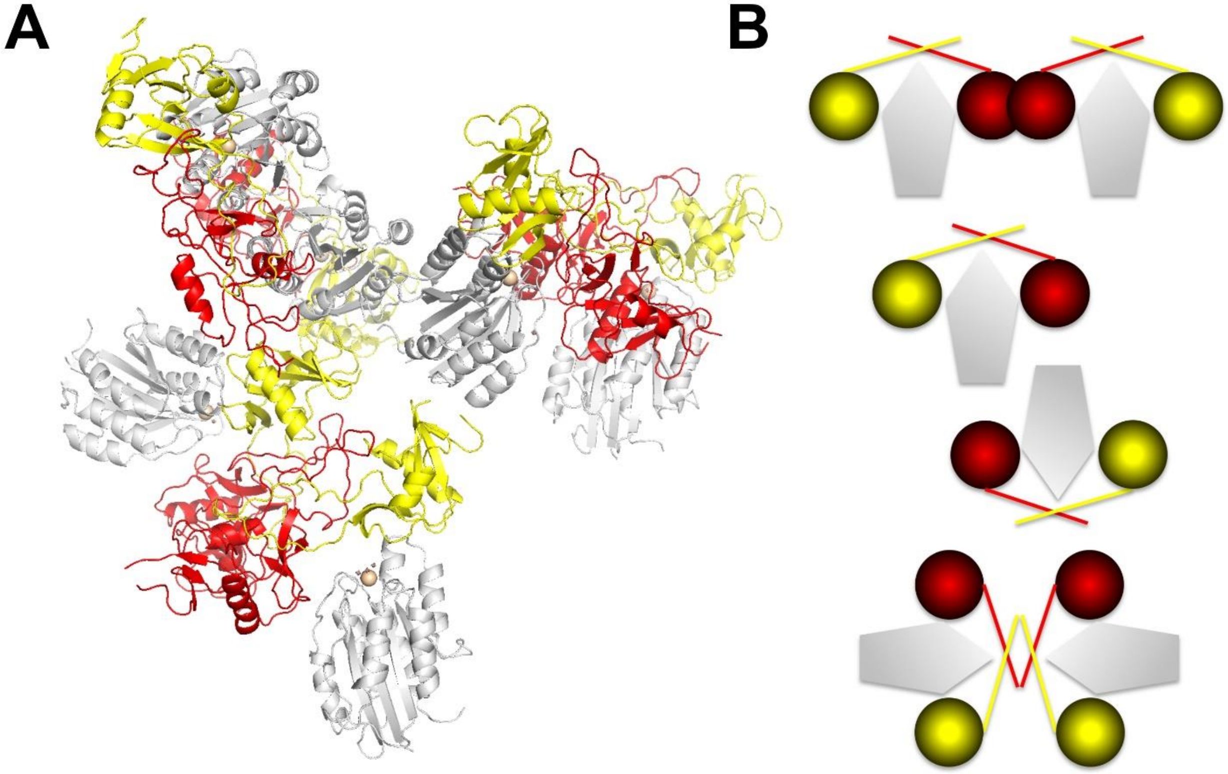


Figure II-1. Asymmetric Unit of the RCγδ-α2A crystal structure.

(**A**) Overall view of the asymmetric unit showing six RCγδ-α2A complexes. Individual α2A domains are shown in grey, with the Mn^{2+} as pink spheres. RCγ subunits are shown in red, whereas RCδ subunits are in yellow. (**B**) The different heterotrimeric assemblies can be subcategorized in three different interaction modii. Domain-domain contacts are mediated either via the core segment of the CLRP fold of RCγ (top), the distal end of the α2A domain (middle) or the index finger loop segments (bottom). Remarkably, the overall r.m.s.d. in Cα positions for all individual subdomains is 1.1Å demonstrating that the different RCγδ−α2A complexes are identical.

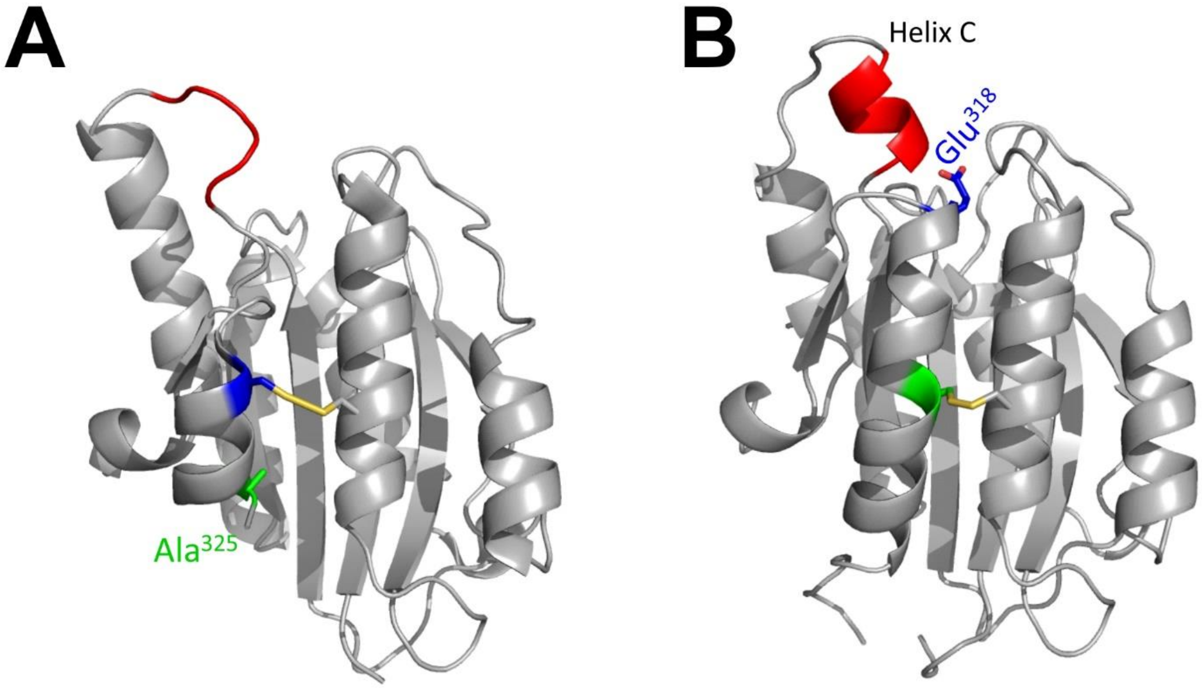


Figure II-2. Molecular model of the disulfide-locked conformation mutants of α2A domain.

By introducing disulfide bridges at the respective sites, helices 1 and 7 were fixed towards each other. Using this approach, the the α2A domain is stabilized in either the “open” or “closed” conformation. (**A**) Model of K168C-E318C representing the open conformation. (**B**) Model of K168C-A325C showing the conformation. Residues involved in the formation of helix C are in red. To highlight the difference between the two conformations, amino acid residue positions 318 and 325 are coloured blue and green, respectively. Structures were modeled with Pymol using the pdb data sets of α2A domain in its “open” (1DZI) and “closed” (1AOX) conformation.

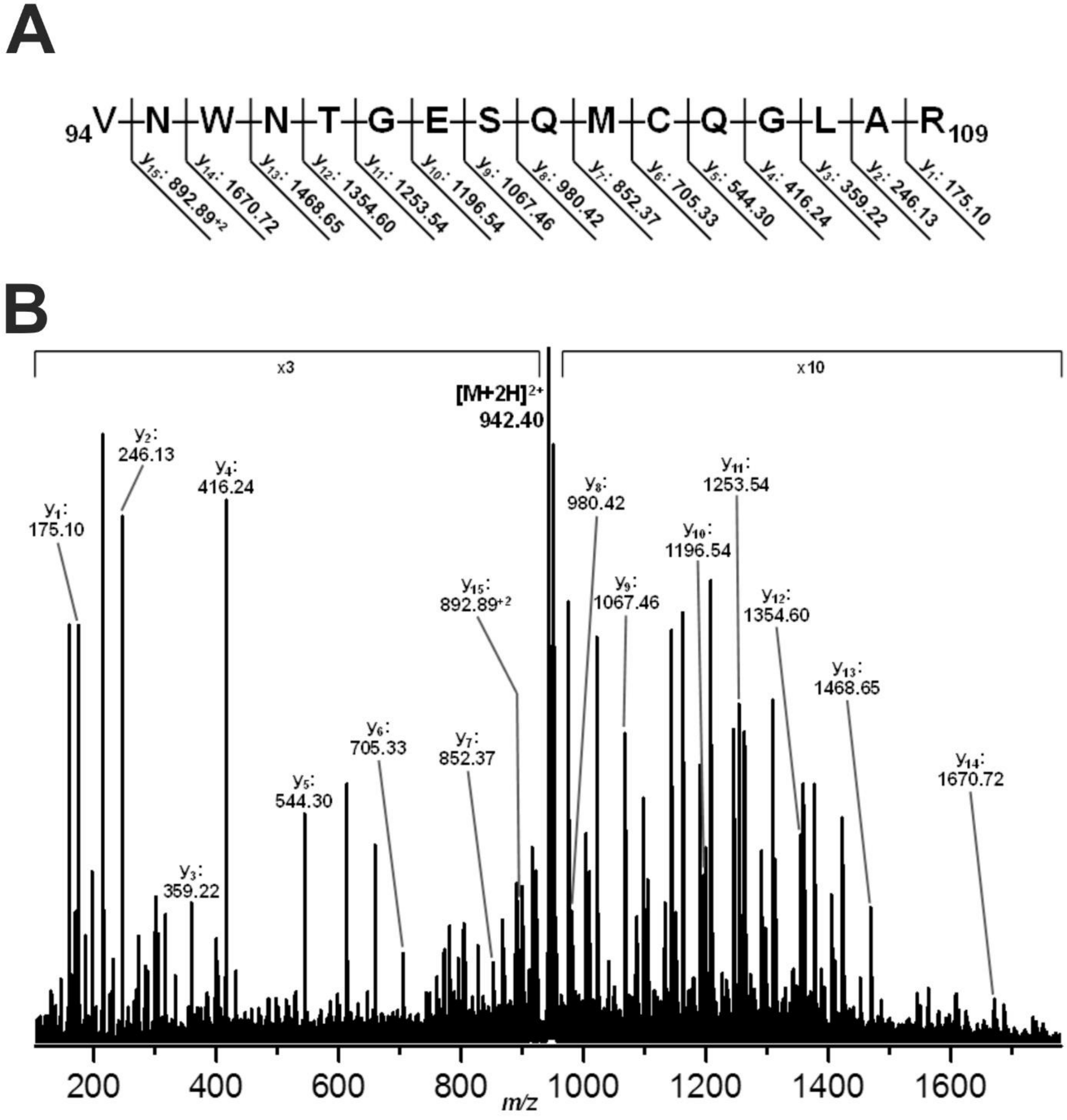
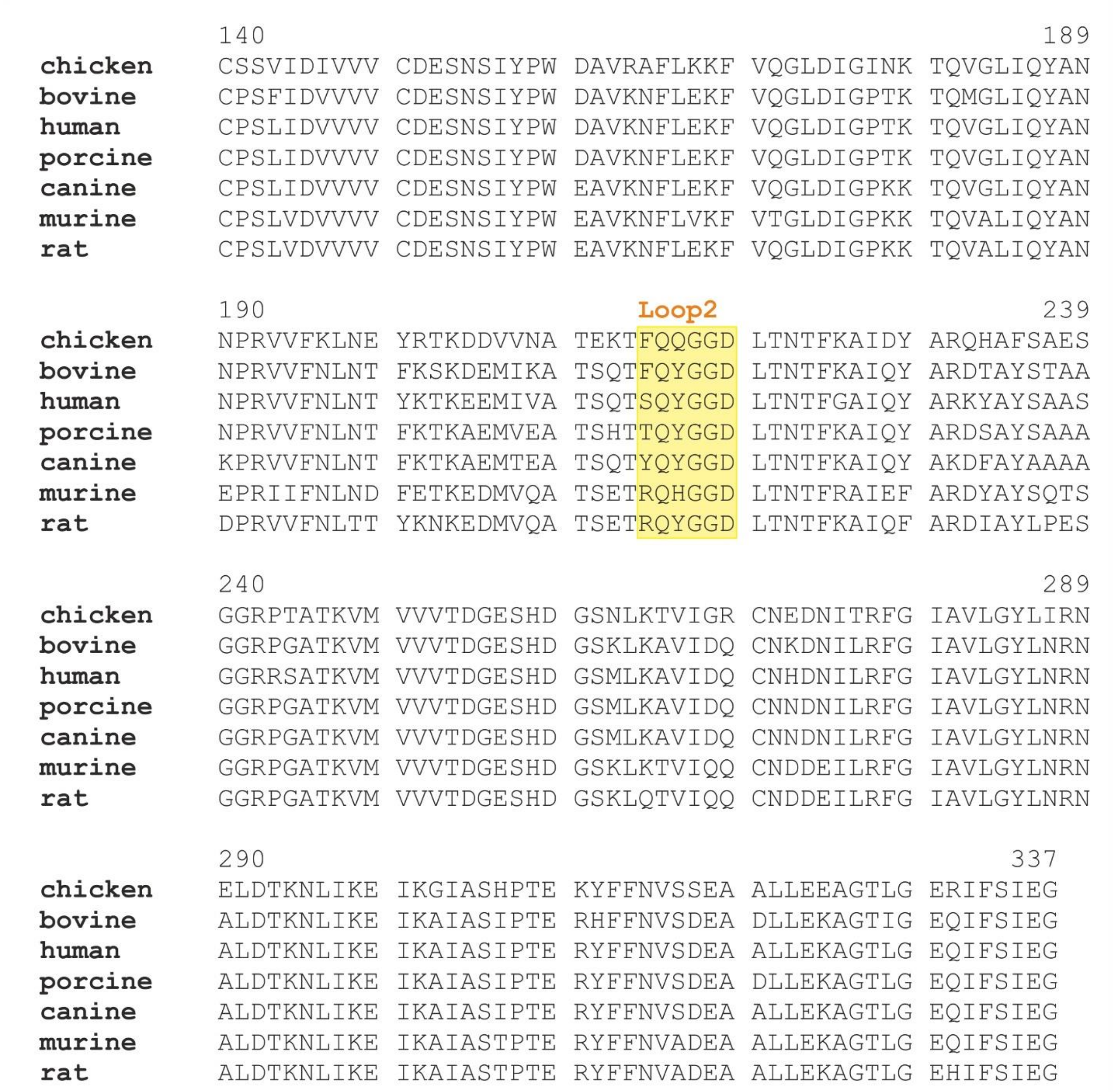


Figure II-3. Identification of the IIIG5 epitope within the RCγ chain.

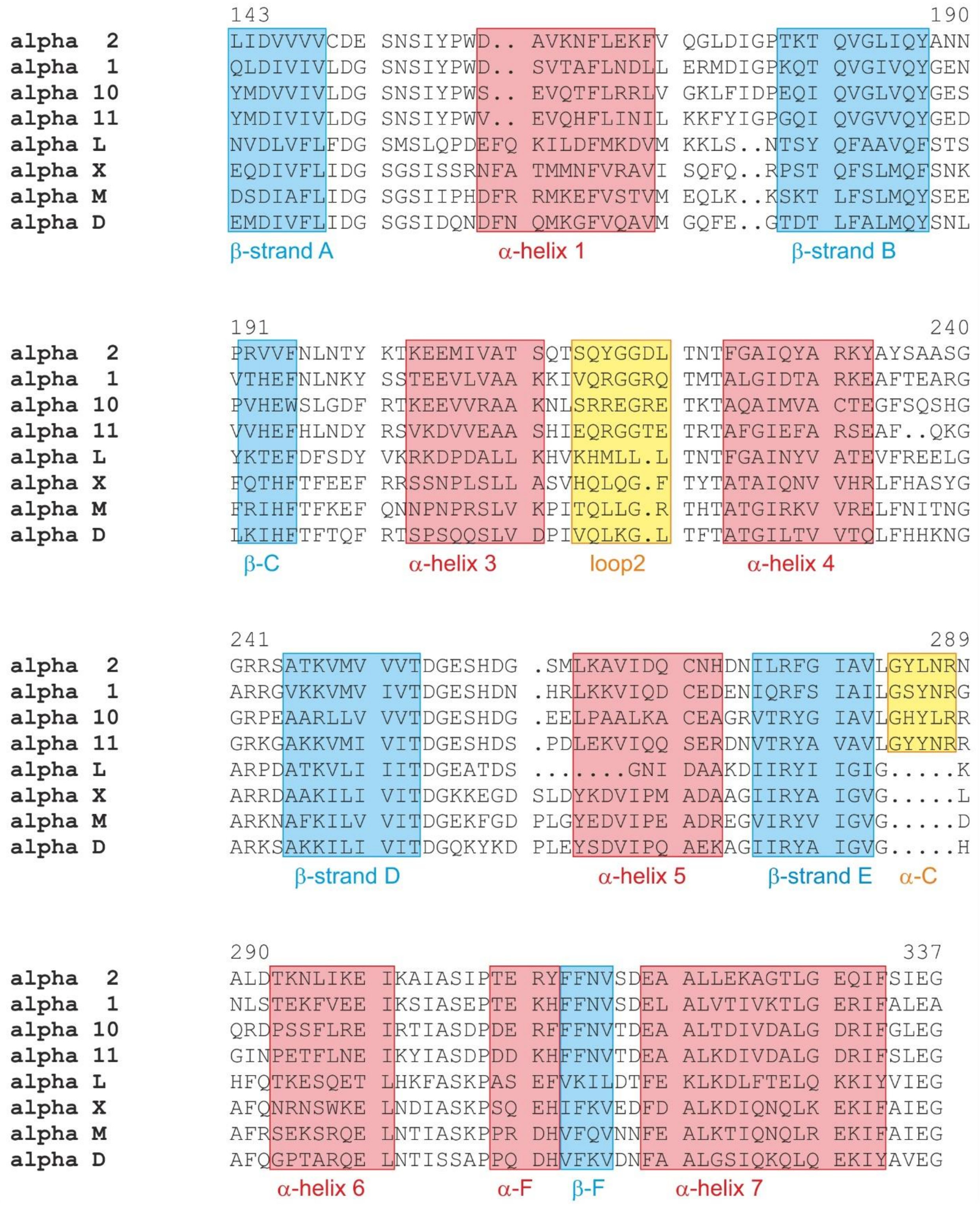
(**A**) Fragmentation scheme for the tryptic fragment of the RCγ subunit containing the IIIG5 epitope. (**B)** NanoESI fragment ion spectrum of the RCγ peptide containing the IIIG5 epitope. It was obtained from a CID experiment on the ion mobility-separated doubly charged peptide precursor ions at m/z 942.40. The labelled peaks correspond to the fragmentation ions of this epitope peptide, as shown in (**A**).



```
            140                                                    189
chicken     CSSVIDIVVV CDESNSIYPW DAVRAFLKKF VQGLDIGINK TQVGLIQYAN
bovine      CPSFIDVVVV CDESNSIYPW DAVKNFLEKF VQGLDIGPTK TQMGLIQYAN
human       CPSLIDVVVV CDESNSIYPW DAVKNFLEKF VQGLDIGPTK TQVGLIQYAN
porcine     CPSLIDVVVV CDESNSIYPW DAVKNFLEKF VQGLDIGPTK TQVGLIQYAN
canine      CPSLIDVVVV CDESNSIYPW EAVKNFLEKF VQGLDIGPKK TQVGLIQYAN
murine      CPSLVDVVVV CDESNSIYPW EAVKNFLVKF VTGLDIGPKK TQVALIQYAN
rat         CPSLVDVVVV CDESNSIYPW EAVKNFLEKF VQGLDIGPKK TQVALIQYAN

            190                          Loop2                     239
chicken     NPRVVFKLNE YRTKDDVVNA TEKTFQQGGD LTNTFKAIDY ARQHAFSAES
bovine      NPRVVFNLNT FKSKDEMIKA TSQTFQYGGD LTNTFKAIQY ARDTAYSTAA
human       NPRVVFNLNT YKTKEEMIVA TSQTSQYGGD LTNTFGAIQY ARKYAYSAAS
porcine     NPRVVFNLNT FKTKAEMVEA TSHTTQYGGD LTNTFKAIQY ARDSAYSAAA
canine      KPRVVFNLNT FKTKAEMTEA TSQTYQYGGD LTNTFKAIQY AKDFAYAAAA
murine      EPRIIFNLND FETKEDMVQA TSETRQHGGD LTNTFRAIEF ARDYAYSQTS
rat         DPRVVFNLTT YKNKEDMVQA TSETRQYGGD LTNTFKAIQF ARDIAYLPES

            240                                                    289
chicken     GGRPTATKVM VVVTDGESHD GSNLKTVIGR CNEDNITRFG IAVLGYLIRN
bovine      GGRPGATKVM VVVTDGESHD GSKLKAVIDQ CNKDNILRFG IAVLGYLNRN
human       GGRRSATKVM VVVTDGESHD GSMLKAVIDQ CNHDNILRFG IAVLGYLNRN
porcine     GGRPGATKVM VVVTDGESHD GSMLKAVIDQ CNNDNILRFG IAVLGYLNRN
canine      GGRPGATKVM VVVTDGESHD GSMLKAVIDQ CNNDNILRFG IAVLGYLNRN
murine      GGRPGATKVM VVVTDGESHD GSKLKTVIQQ CNDDEILRFG IAVLGYLNRN
rat         GGRPGATKVM VVVTDGESHD GSKLQTVIQQ CNDDEILRFG IAVLGYLNRN

            290                                                337
chicken     ELDTKNLIKE IKGIASHPTE KYFFNVSSEA ALLEEAGTLG ERIFSIEG
bovine      ALDTKNLIKE IKAIASIPTE RHFFNVSDEA DLLEKAGTIG EQIFSIEG
human       ALDTKNLIKE IKAIASIPTE RYFFNVSDEA ALLEKAGTLG EQIFSIEG
porcine     ALDTKNLIKE IKAIASIPTE RYFFNVSDEA DLLEKAGTLG EQIFSIEG
canine      ALDTKNLIKE IKAIASIPTE RYFFNVSDEA ALLEKAGTLG EQIFSIEG
murine      ALDTKNLIKE IKAIASTPTE RYFFNVADEA ALLEKAGTLG EQIFSIEG
rat         ALDTKNLIKE IKAIASTPTE RYFFNVADEA ALLEKAGTLG EHIFSIEG
```

Figure II-4. Alignment of integrin α2A domains from different species.

Sequence comparison of the integrin α2 A-domain from different vertebrate species. The loop 2 sequence S^{214}QYGGD is highlighted in yellow and shows a high degree of homology between different species. Multiple sequence alignment was carried out with Clustal Omega Software from EMBL-EBI.



```
                143                                                 190
alpha  2        LIDVVVVCDE SNSIYPWD.. AVKNFLEKFV QGLDIGPTKT QVGLIQYANN
alpha  1        QLDIVIVLDG SNSIYPWD.. SVTAFLNDLL ERMDIGPKQT QVGIVQYGEN
alpha 10        YMDVVIVLDG SNSIYPWS.. EVQTFLRRLV GKLFIDPEQI QVGLVQYGES
alpha 11        YMDIVIVLDG SNSIYPWV.. EVQHFLINIL KKFYIGPGQI QVGVVQYGED
alpha L         NVDLVFLFDG SMSLQPDEFQ KILDFMKDVM KKLS..NTSY QFAAVQFSTS
alpha X         EQDIVFLIDG SGSISSRNFA TMMNFVRAVI SQFQ..RPST QFSLMQFSNK
alpha M         DSDIAFLIDG SGSIIPHDFR RMKEFVSTVM EQLK..KSKT LFSLMQYSEE
alpha D         EMDIVFLIDG SGSIDQNDFN QMKGFVQAVM GQFE..GTDT LFALMQYSNL

                191                                                 240
alpha  2        PRVVFNLNTY KTKEEMIVAT SQTSQYGGDL TNTFGAIQYA RKYAYSAASG
alpha  1        VTHEFNLNKY SSTEEVLVAA KKIVQRGGRQ TMTALGIDTA RKEAFTEARG
alpha 10        PVHEWSLGDF RTKEEVVRAA KNLSRREGRE TKTAQAIMVA CTEGFSQSHG
alpha 11        VVHEFHLNDY RSVKDVVEAA SHIEQRGGTE TRTAFGIEFA RSEAF..QKG
alpha L         YKTEFDFSDY VKRKDPDALL KHVKHMLL.L TNTFGAINYV ATEVFREELG
alpha X         FQTHFTFEEF RRSSNPLSLL ASVHQLQG.F TYTATAIQNV VHRLFHASYG
alpha M         FRIHFTFKEF QNNPNPRSLV KPITQLLG.R THTATGIRKV VRELFNITNG
alpha D         LKIHFTFTQF RTSPSQQSLV DPIVQLKG.L TFTATGILTV VTQLFHHKNG

                241                                                 289
alpha  2        GRRSATKVMV VVTDGESHDG .SMLKAVIDQ CNHDNILRFG IAVLGYLNRN
alpha  1        ARRGVKKVMV IVTDGESHDN .HRLKKVIQD CEDENIQRFS IAILGSYNRG
alpha 10        GRPEAARLLV VVTDGESHDG .EELPAALKA CEAGRVTRYG IAVLGHYLRR
alpha 11        GRKGAKKVMI VITDGESHDS .PDLEKVIQQ SERDNVTRYA VAVLGYYNRR
alpha L         ARPDATKVLI IITDGEATDS .......GNI DAAKDIIRYI IGIG.....K
alpha X         ARRDAAKILI VITDGKKEGD SLDYKDVIPM ADAAGIIRYA IGVG.....L
alpha M         ARKNAFKILV VITDGEKFGD PLGYEDVIPE ADREGVIRYV IGVG.....D
alpha D         ARKSAKKILI VITDGQKYKD PLEYSDVIPQ AEKAGIIRYA IGVG.....H

                290                                                337
alpha  2        ALDTKNLIKE IKAIASIPTE RYFFNVSDEA ALLEKAGTLG EQIFSIEG
alpha  1        NLSTEKFVEE IKSIASEPTE KHFFNVSDEL ALVTIVKTLG ERIFALEA
alpha 10        QRDPSSFLRE IRTIASDPDE RFFFNVTDEA ALTDIVDALG DRIFGLEG
alpha 11        GINPETFLNE IKYIASDPDD KHFFNVTDEA ALKDIVDALG DRIFSLEG
alpha L         HFQTKESQET LHKFASKPAS EFVKILDTFE KLKDLFTELQ KKIYVIEG
alpha X         AFQNRNSWKE LNDIASKPSQ EHIFKVEDFD ALKDIQNQLK EKIFAIEG
alpha M         AFRSEKSRQE LNTIASKPPR DHVFQVNNFE ALKTIQNQLR EKIFAIEG
alpha D         AFQGPTARQE LNTISSAPPQ DHVFKVDNFA ALGSIQKQLQ EKIYAVEG
```

Figure II-5. Alignment of A-domain of different human integrin α-chains.

A comparison of A-domains from different human integrin α subunits. Integrin alpha subunits 1, 2, 10, and 11 belong to the subset of collagen binding integrins. They possess the characteristic helix C (yellow box, labelled α-C), which is absent in the A-domain of the leukocyte β2 integrins with their alpha subunits L, X, M, and D. Helix C of the integrin α2 subunit is the primary binding site for RCγδ and is only present in the "closed" conformation of its A domain. The secondary RC contact site of α2A is located within the loop 2 sequence S^{214}QYGGD, (yellow box, labelled loop 2) and is specific to the integrin α2 chain. The secondary structure elements are indicated by the red (α-helices) and the blue (β-strands) boxes, respectively. The residue numbering refers to the integrin α2 sequence. Multiple sequence alignment was carried out with Clustal Omega Software from EMBL-EBI.

Appendix III: Additional Supporting Information for Archaea S-layer nanotube from a "black smoker" in complex with cyclo-octasulfur S_8 rings

Acknowledgements

The presented work was funded by the Discovery Grant (RGPIN-342077-2012) program, the Strategic Partnership Program (STGP-479210-2015) and the Research Tool and Infrastructure support (345517-07) from NSERC. JS is a Canada Research Chair in Structural Biology. We thank the Canadian Light Source for beamtime.

Figures

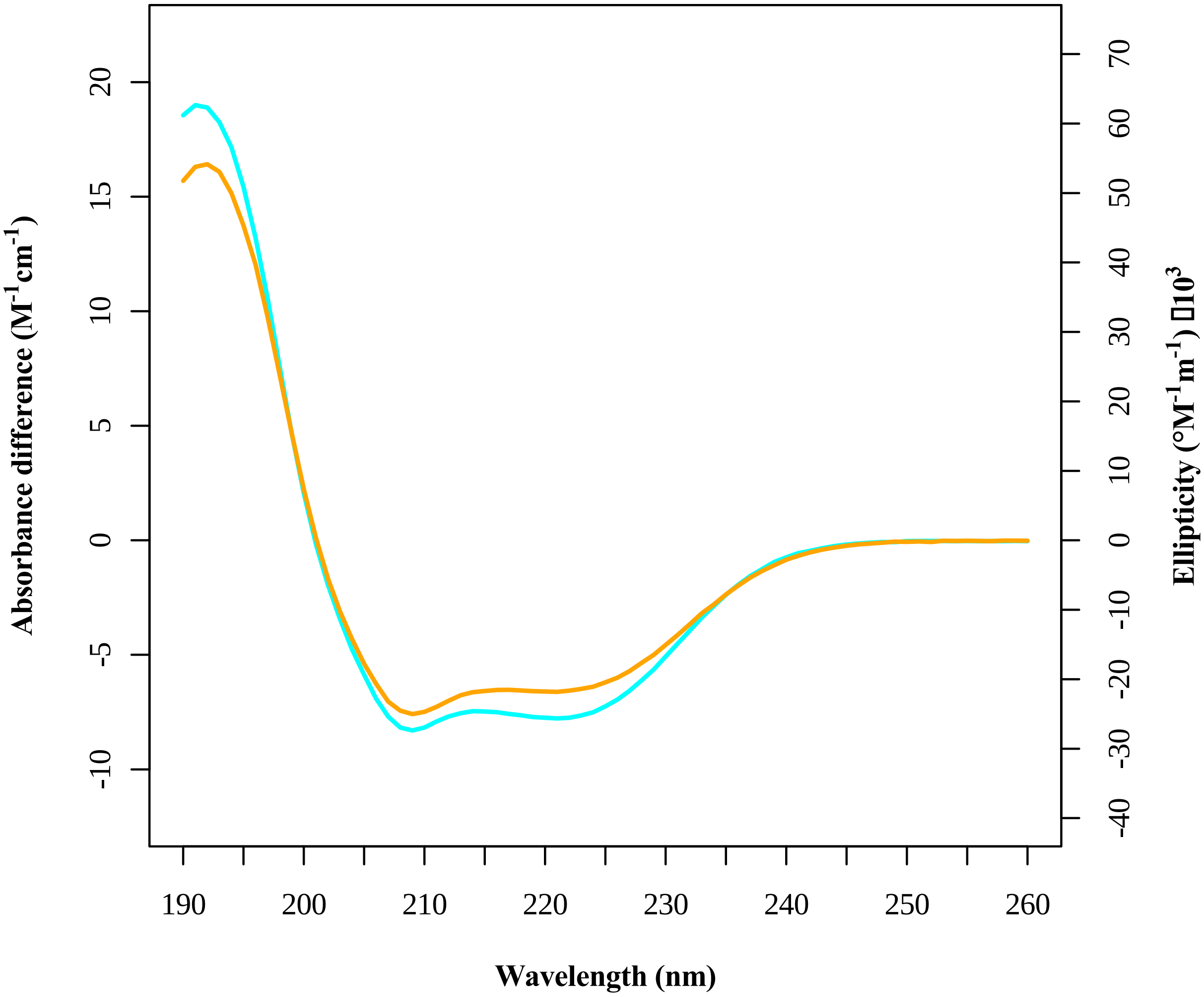


Figure III-1. CD experiments.

CD measurements of RHCC-NT at 20°C (cyan) and 81°C (orange) demonstrate that the homo-tetrameric assembly of RHCC maintains significant α-helicity at elevated temperatures.

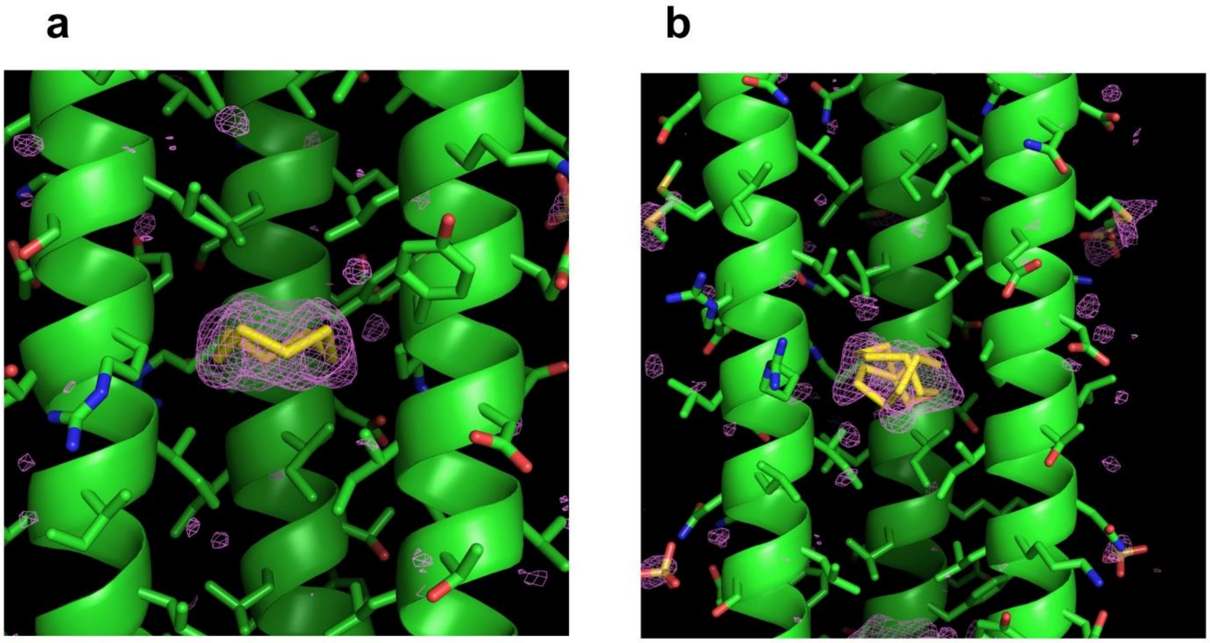


Figure III-2. S-SAD electron density.

3 Sigma anomalous difference map at $\lambda = 1.77123$Å. The map was calculated using experimental ΔF_{ano} and φ_{calc} from apoRHCC (pdb 1FE6). **a**, View inside cavity 2 with a single S8-crown. **b**, View of the tumbling S_8 crown in cavity 3.

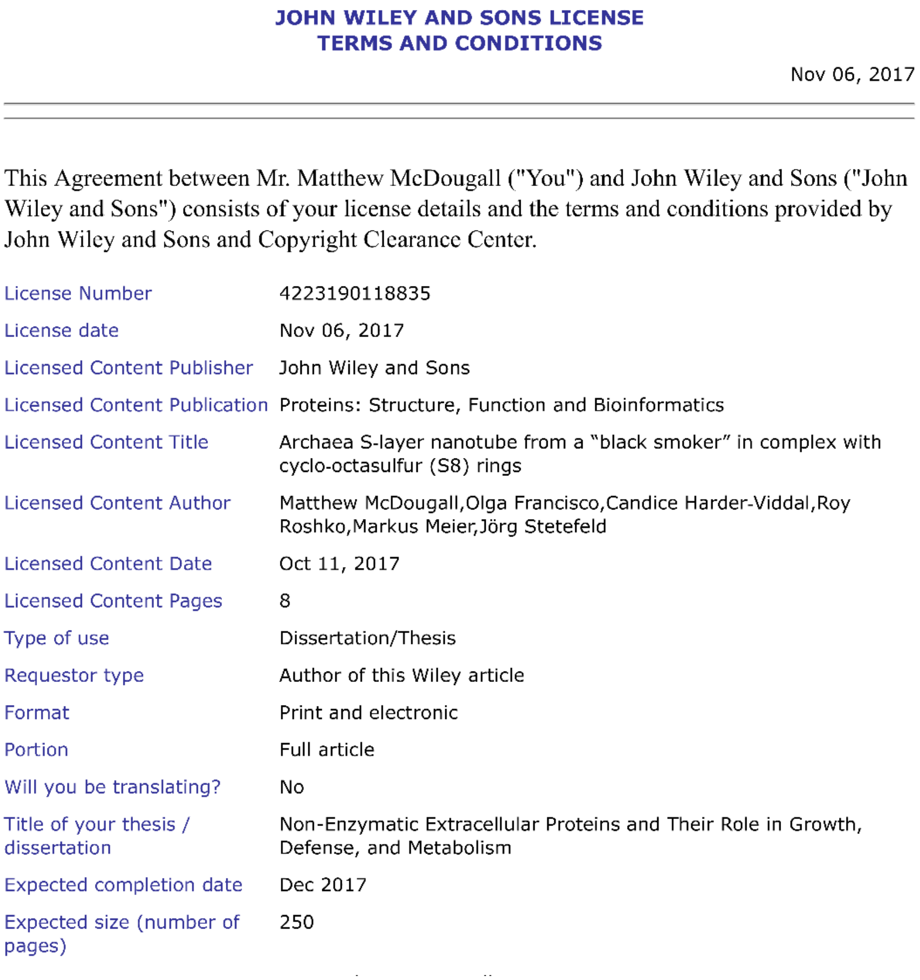
$$\left[\left(LP\right)_{sol}\right]_{V_{MD}} \xrightarrow{\Delta G_1} \left[\left(\boxed{L}P\right)_{sol}\right]_{V_{MD}} \xrightarrow{\Delta G_2} \left[\left(\boxed{L}_{gas}P\right)_{sol}\right]_{V_{MD}} \xrightarrow{\Delta G_3} \left[\left(L\right)_{gas}\right]_{V_0}$$

$$\left[\left(L\right)_{sol}\right]_{V_{MD}} \xrightarrow{\Delta G_4} \left[\boxed{L}_{sol}\right]_{V_{MD}} \xrightarrow{\Delta G_5} \left[\boxed{L}_{gas}\right]_{V_{MD}}$$

Figure III-3. Thermodynamic Paths for Free Energy Calculations.

A schematic diagram showing the two thermodynamic paths used to calculate the absolute binding free energy for the transfer of a ligand from the solvent bath to cavity 2 of RHCC tetrabrachion. The ligand is either the sulfur S_8 crown or the $[H_2O]_9$ water cluster. The rectangle represents the FBHW restraint.

Right to Reproduce Licence



JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Nov 06, 2017

This Agreement between Mr. Matthew McDougall ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4223190118835
License date	Nov 06, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Proteins: Structure, Function and Bioinformatics
Licensed Content Title	Archaea S-layer nanotube from a “black smoker” in complex with cyclo-octasulfur (S8) rings
Licensed Content Author	Matthew McDougall,Olga Francisco,Candice Harder-Viddal,Roy Roshko,Markus Meier,Jörg Stetefeld
Licensed Content Date	Oct 11, 2017
Licensed Content Pages	8
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Non-Enzymatic Extracellular Proteins and Their Role in Growth, Defense, and Metabolism
Expected completion date	Dec 2017
Expected size (number of pages)	250