Structure-Function Analyses of

Two-Component Signaling in *Arabidopsis Thaliana* and

Cellular Attachment Mechanisms of Adenovirus Type 37

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Johannes Bauer

aus VS-Villingen

Tübingen

2013

Table of Contents

Abbreviations

Ad	Adenovirus
AHK	Arabidopsis histidine kinase
AHP	Arabidopsis histidine-containing phosphotransfer protein
AIDS	Acquired immune deficiency syndrome
ARR	Arabidopsis response regulator
Asp	Aspartic acid
ASU	Asymmetric unit
A. thaliana	*Arabidopsis thaliana*
BiFC	Bimolecular fluorescence complementation
CA	Catalytic domain
CAR	Coxsackie- and adenovirus receptor
CD	Cluster of differentiation
CK	Cytokinin
CKI	Cytokinin independent
CRM	Chromosome region maintenance
CTR	Constitutive triple response
Da	Dalton
DHP	Dimerization- and histidine-containing phosphotransfer domain
DNA	Desoxyribonucleic acid
dsDNA	double stranded DNA
DSG-2	Desmoglein-2
E. coli	*Escherichia coli*
e. g.	example given
EIL	EIN3-like
EIN	Ethylene insensitive
EKC	Epidemic keratoconjunctivitis
EM	Electron microscopy
ER	Endoplasmatic reticulum
ERS	Ethylene response sensor
ETR	Ethylene resistant

GAF	c**G**MP-specific phosphodiesterases, **a**denylyl cyclases and **F**hlA
Gly	Glycin
HCE cells	Human corneal epithelial cells
His	Histidine
HIV	Human immunodeficiency virus
HK	Histidine kinase
HSA	Human serum albumin
Hsp	Heat shock protein
IC_{50}	Half maximal inhibitory concentration
i. e.	id est
bp	base pair
Lys	Lysine
MAP	Mitogen-activated protein
MSP	Multi-step phosphorelay
MTOC	Microtubule-organizing center
$NMSO_3$	Sodium[2,2-bis(docosyloxymethyl)propyl-5-acetoamido-3,5-dideoxy-4,7,8,9-tetra-*O*-(sodiumoxysulfonyl)-D-glycero-α-D-galacto-2-nonulo-pyranosid]onate
NPC	Nuclear pore complex
PAS	**P**eriod circadian protein, **a**ryl hydrocarbon receptor nuclear translocator protein, **s**ingle-minded protein
Pro	Proline
PRR	Pseudo response regulator
RAF	Rapidly accelerated fibrosarcoma
RAN	Responsive to antagonist
RD	Receiver domain
ROS	Reactive oxygen species
R. m. s.	Root mean square
RNA	Ribonucleic acid
RR	Response regulator
R. sphaeroides	*Rhodobacter sphaeroides*
RSV	Respiratory syncytial virus

RTE	Reversion to ethylene sensitivity
Ser	Serine
SPR	Surface plasmon resonance
TCS	Two-component system
Thr	Threonine
Tyr	Tyrosine
Val	Valine

1 Introduction Part I: Plant Two-Component Signaling

1.1 Signal transduction in plants

Plants live in very different habitats all over the world. Therefore, they are exposed to a wide variety of environmental challenges, such as extreme temperatures, light and darkness, drought and flooding, lack of nutrients, and salinity, but they also have to withstand parasites, pathogens and herbivores. This is particularly remarkable as plants are not mobile and therefore cannot escape their environment. Thus, in order to survive it is absolutely essential for plants to develop strategies that allow for optimal adaption to their environmental challenges. During evolution, plants often found many different ways to cope with such environmental challenges, which is most apparently expressed in the extreme diversity of plant size and morphology. However, on a cellular basis, the involved molecular mechanisms that translate environmental signals into intracellular responses share fundamental similarities among plants and are often related to those of other organisms. In principle, plants must be able to sense environmental stimuli, transmit these signals within and across cells, and finally respond appropriately to the input signals.

1.2 The two-component system (TCS) in *Arabidopsis thaliana*

1.2.1 Principles of two-component signaling

Two-component signaling systems serve as the foundation of intracellular communication in many biological systems. They are wide-spread in prokaryotes but TCS-elements have also been identified in many eukaryotes including plants. Interestingly, however, they are absent in animals. The most basic version of the TCS comprises two proteins, a histidine kinase (HK) and a response regulator (RR) (Figure 1A). The HK senses a certain stimulus via its sensor domain, leading to autophosphorylation of a conserved histidine residue within the dimerization- and histidine-containing phosphotransfer domain of the HK (HK_{DHP}). The phosphoryl group is then intermolecularly transferred onto a conserved aspartic acid residue in the RR, which then primarily regulates target genes, or - to a lesser extent - target proteins (Mascher et al., 2006).

In the year 2000, the genomic analysis of the plant *Arabidopsis thaliana* led to the identification of about 40 proteins that function in a more complex version of the TCS, known as the multi-step phosphorelay (MSP) (Arabidopsis Genome Initiative, 2000). Since

then, the MSP is recognized as a fundamental signaling pathway in plants. Compared to the basic TCS, the MSP contains additional proteins and domains (Figure 1B). Thus, the MSP system in *A. thaliana* comprises three protein types: a sensor HK, a histidine-containing phosphotransfer protein (AHP) and a response regulator (ARR). The HK perceives endo- or exogenous stimuli via its sensor domain, leading to autophosphorylation of a conserved histidine residue within the cytoplasmic HK_{DHP} domain. The phosphoryl group is then relayed onto a conserved aspartic acid residue in the receiver domain of the kinase (HK_{RD}). Next, the phosphoryl group is transferred from HK_{RD} to an invariant histidine residue in the downstream AHP, which finally delivers the phosphoryl group to a conserved aspartic acid residue in the cognate ARR. Activated ARRs then regulate their respective target proteins or act as transcription factors (Grefen and Harter, 2004; Horak et al., 2011; Laub and Goulian, 2007; Schaller et al., 2008).
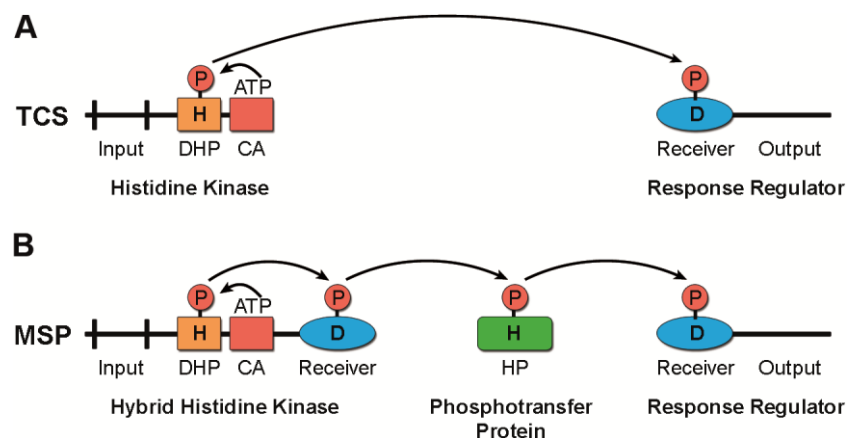

Figure 1. Schematic overview of His-Asp phosphorelay systems; taken from Bauer et al., 2013. **(A)** Overview of phosphorelay signaling in two-component systems (TCS). Signal perception induces ATP-binding to the catalytic domain (CA), followed by phosphorylation of a conserved histidine residue in the dimerization and phosphotransfer domain (DHP) of the kinase. Finally, the phosphoryl group (P) is relayed onto a conserved aspartic acid residue in the receiver domain of the cognate response regulator. **(B)** Overview of phosphorelay signaling in multi-step phosphorelay systems (MSP). MSP signaling involves a sensor histidine kinase containing a fused receiver domain and an additional HP protein.

1.2.2   MSP elements in *A. thaliana*

The *A. thaliana* genome encodes eight canonical HKs (AHK1-5, ETR1, ERS1, and CKI1), five canonical AHPs (AHP1-5), and 23 canonical ARRs (ARR1-22, ARR24) (Figure 2). In addition, there are nine pseudo-HKs, one pseudo-AHP (AHP6) and nine pseudo-ARRs (PRR1-9) that carry mutations at the critical histidine or aspartic acid residues, respectively (Grefen and Harter, 2004; Schaller et al., 2008).

Within the *A. thaliana* sensor HK family, ETR1 and ERS1 are members of the ethylene receptor family, whereas AHK2-4 are grouped into the cytokinin receptor family (Grefen and Harter, 2004). The remaining three HKs AHK1, CKI1, and AHK5 (also known as CKI2) do not belong to a subfamily. The AHP proteins serve as phosphoryl carrier between sensor HKs and ARRs, and importantly they can transfer the signal from the cytoplasm into the nucleus, the location of many ARRs. The ARRs are subclassified into types A, B, and C based on phylogenetic and functional analysis and domain architecture (Schaller et al., 2008). B-type ARRs (ARR1, 2, 10-14, 18-21) directly act as transcription factors, thereby also regulating A-type ARRs (ARR3-9, ARR15-17) which in turn act on their respective target proteins. In contrast, C-type ARRs (ARR22 and ARR24) are poorly understood. They share with A-type ARRs the lack of the Myb-like DNA-binding domain (Schaller et al., 2008), although they are not closely related to A-type ARRs based on sequence and phylogenetic analysis (Gattolin et al., 2006; Schaller et al., 2008).
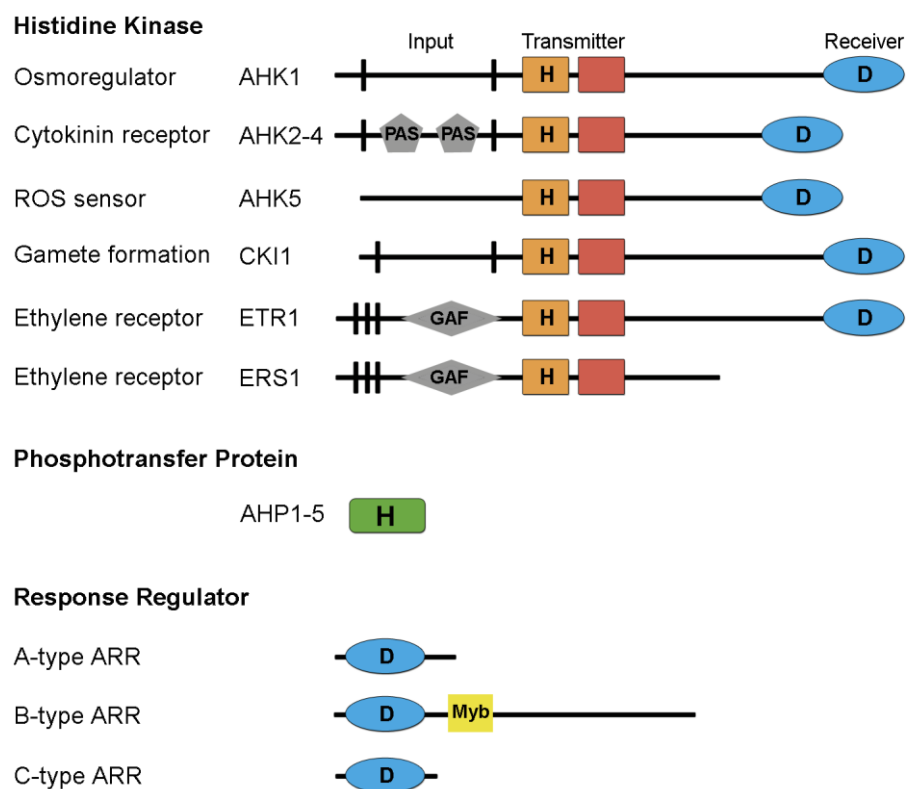

Figure 2. Domain organization of TCS-elements in *A. thaliana*; adapted from Schaller et al., 2008.

Ethylene receptor family

The gaseous hormone ethylene effects and regulates various processes within the plant life cycle. These include seed germination, seedling growth and development, flowering and fruit ripening, organ senescence, and leaf abscission. Additionally, ethylene also contributes to stress and pathogen responses (Abeles et al., 1992). In *A. thaliana*, the ethylene receptors act as negative regulators in phophorelay signaling (Hua and Meyerowitz, 1998), and thus they actively repress the ethylene response in absence of the hormone (Taiz, 2010). Upon ethylene binding, autophosphorylation is inhibited, leading to release of the pathway repression and thus allowing functional ethylene signaling (Bisson and Groth, 2011; Voet-van-Vormizeele and Groth, 2008).

The ethylene receptor family in *A. thaliana* comprises five proteins: ETR1, ERS1, ETR2, ERS2, and EIN4. Based on functional and phylogenetic analyses, the ethylene receptors are grouped into two subfamilies. Subfamily I receptors, comprising ETR1 and ERS1, contain a functional HK domain (Gamble et al., 1998; Moussatche and Klee, 2004). In contrast, subfamily II members (ETR2, ERS2, and EIN4) carry mutations in critical catalytic residues that result in loss of HK activity. Instead, they have evolved to act as serine/threonine kinases (Moussatche and Klee, 2004).

All ethylene receptors share an N-terminal transmembrane domain followed by a GAF-domain and a kinase domain. In addition, ETR1, EIN4, and ETR2 contain a receiver domain fused to the C-terminus of the kinase domain (Voet-van-Vormizeele and Groth, 2008). The GAF domains were found to mediate non-covalent association of ethylene receptors, indicating that the ethylene receptors exist in plants as multimeric clusters (Chen et al., 2010; Gao et al., 2008; Grefen et al., 2008).

The N-terminal transmembrane domain bears the ethylene binding site, which can bind one ethylene molecule per receptor dimer via a copper ion (Rodriguez et al., 1999). Furthermore, the transmembrane domain anchors all ethylene receptors to the ER membrane (Chen et al., 2002; Grefen et al., 2008). In addition, the N-terminus of ETR1 is tightly associated with RTE1, which is thought to regulate the signaling state of the receptor complex (Dong et al., 2010). Interestingly, at least ETR1 also co-localizes to the Golgi apparatus (Dong et al., 2008). This is in-line with experiments indicating that the P-type ATPase RAN1 co-localizes with ETR1 at the Golgi apparatus (Binder et al., 2010; Dong et al., 2008). Thus, RAN1 could provide ETR1 with the essential Cu(I) co-factor for ethylene binding.

Downstream signal transduction from the ethylene receptor family in the ER membrane towards the transcription factor family of EIN3/EIL proteins in the nucleus involves the RAF-like kinase CTR1 and the ER-membrane protein EIN2. Thereby, CTR1 acts as negative regulator on EIN2 in absence of ethylene (Kieber et al., 1993). Ethylene binding first inhibits phosphorylation of the ethylene receptors, which then leads in a second step to inactivation of CTR1 (Bisson and Groth, 2011). This in turn allows ETR1 to tightly interact with EIN2 at the ER membrane (Bisson et al., 2009), and finally to switch on the transcription of ethylene response genes (Figure 3).
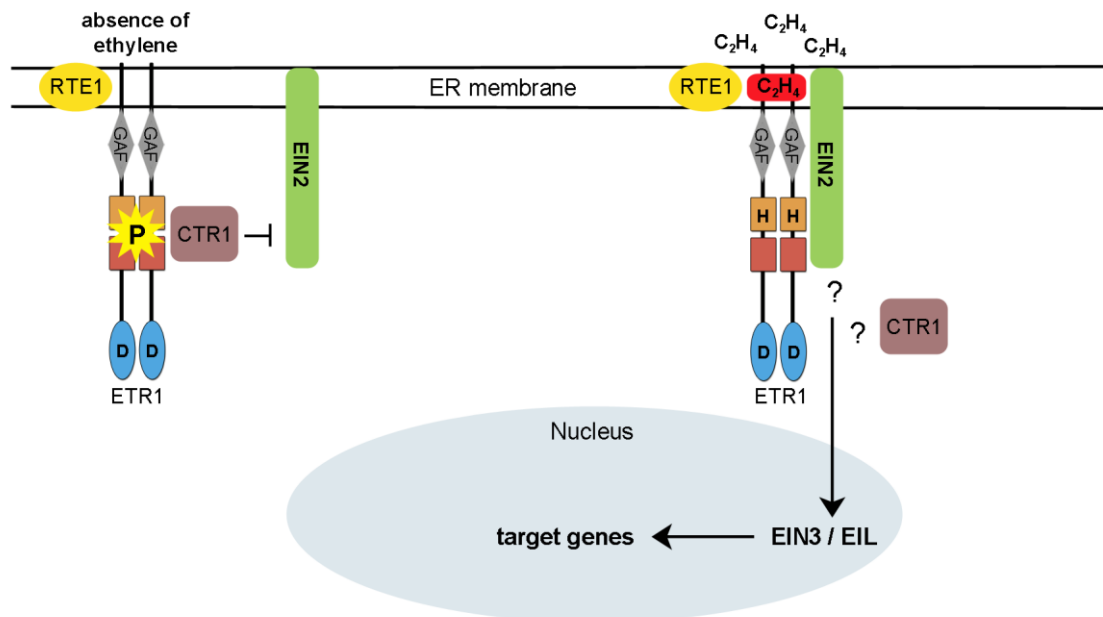
Figure 3. Model of the ethylene signaling pathway in *A. thaliana*. Ethylene binding at the N-terminal transmembrane domain of the ETR1 homodimer inhibits autophosphorylation. Inactivation of CTR1 then allows EIN2 to tightly interact with the ETR1 kinase domain and to activate the family of EIN3/EIL transcription factors.

Cytokinin receptor family

The three HKs AHK2, AHK3, and AHK4 represent the cytokinin (CK) receptor family. Naturally occurring CKs are N^6-substituted adenine derivatives that play a central role in plant cell division. In addition, CKs are also involved in various aspects of plant growth and development (Argueso et al., 2010; Hwang et al., 2012; Kieber et al., 1993). Recently, the structural basis of CK recognition by AHK4 has been defined, revealing a highly adaptable binding pocket (Figure 4) (Hothorn et al., 2011).

Figure 4. Structure of the AHK4 sensor domain in complex with N^6-isopentenyladenine. The AHK4 homodimer is shown with its N-terminal stalk (orange), the membrane-distal PAS domain (dark blue) and its membrane-proximal PAS domain (light blue). One N^6-isopentenyladenine molecule is bound per protomer; adapted from Hothorn et al., 2011.
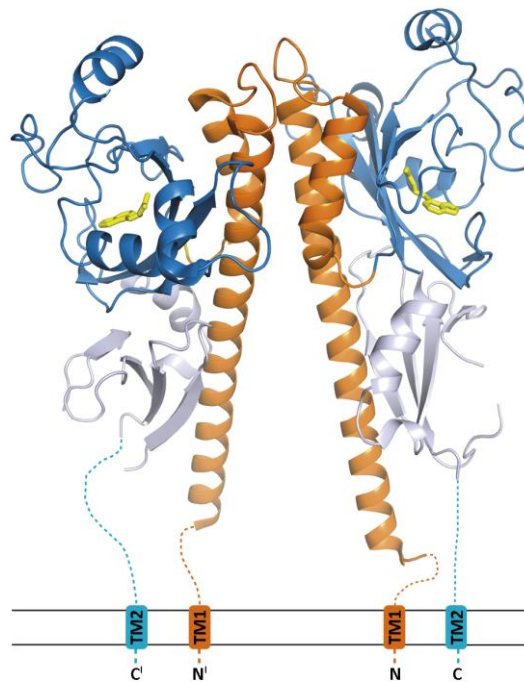
CKs bind to the membrane-distal of two PAS-like domains in the N-terminal sensor domain of AHK4 (Hothorn et al., 2011) and thus initiate the MSP signaling cascade towards B-type ARRs. The nuclear B-type ARRs act as DNA-binding transcription factors and directly promote the expression of nuclear A-type ARRs as primary CK target genes and negative feedback regulators (Hwang et al., 2012). Importantly, there is emerging evidence for a complex network of crosstalk by CK, ethylene and other plant hormones such as auxin, gibberellins, and brassinosteroids, dependent on the environmental condition and the plant's developmental stage (Grefen et al., 2008). Consistently, both CK and ethylene receptors were found to localize to the ER (Caesar et al., 2011; Chen et al., 2002; Grefen et al., 2008; Wulfetange et al., 2011), indicating that this is a site of hormonal crosstalk (Caesar et al., 2011).

AHK1

AHK1 acts as positive regulator in response to drought, high salt, and osmotic stress (Chen et al., 2009; Tran et al., 2007; Wohlbach et al., 2008). Overexpression of AHK1 results in increased tolerance to drought, high salinity, and osmotic stress in the growing plant (Chen

et al., 2009). AHK1-based stress tolerance keeps the physiological cell membrane permeability intact, thereby preventing ion leakage that would lead to plant death (Chen et al., 2009).

CKI1

The sensor HK CKI1 was found to be essential for female gametophyte development and vegetative growth (Deng et al., 2010; Hejatko et al., 2003; Pischke et al., 2002). Recently, the receiver domain of CKI1 (CKI1$_{RD}$) has been structurally characterized (Figure 5) (Pekarova et al., 2011). However, the physiological signal that triggers CKI1 is still unknown.
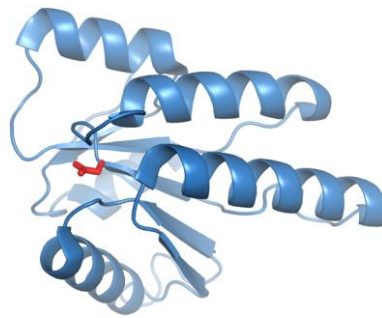
Figure 5. Crystal structure of CKI1$_{RD}$. The conserved aspartic acid for phosphoryltransfer is shown in red stick representation.

AHK5

Lately, the sensor HK AHK5 was found to positively regulate salt sensitivity and to contribute to resistance to the bacterium *Pseudomonas syringae* and the fungal pathogen *Botrytis cinerea* (Pham et al., 2012). Additionally, AHK5 mediates stomatal responses to exogenous and endogenous signals such as flagellin, which alter both the intracellular level of reactive oxygen species (ROS) and redox homeostasis (Desikan et al., 2008). The suggested role of AHK5 as an intracellular ROS sensor is in line with the predicted lack of a transmembrane domain, which renders AHK5 unique among all HKs in *A. thaliana* (Grefen and Harter, 2004).

1.3 Evolution of plant TCS signaling

TCS genes are found in all three kingdoms of life, including prokaryotes, eukaryotes, and archaea (Perraud et al., 1999). However, they are most abundant in the genomes of gram-negative bacteria and cyanobacteria (Capra and Laub, 2012). The modular nature of HKs and RRs separates signal input, phosphotransfer, and output response. This allowed bacteria to

dramatically expand and diversify their signaling capabilities (Capra and Laub, 2012). Interestingly, at least in bacteria the number of signaling elements in general and of TCS elements in particular corresponds to the complexity of the bacterial lifestyle and its capability to face and resist environmental challenges (Alm et al., 2006; Schaller et al., 2011). Plants most likely acquired signal transduction systems by incorporation of endosymbiotic cyanobacteria that finally developed into chloroplasts. This endosymbiotic evolution was accompanied by a drastic reduction of the chloroplast genome. Thereby, genes were either lost or transferred into the nucleus of plant cells (Abdallah et al., 2000; Martin et al., 1998). In total, such lateral gene transfer events are thought to be responsible for the acquisition of approximately 18 % of the Arabidopsis proteome, thereby including at least 189 plant signaling proteins (Martin et al., 2002). Interestingly, proteins of cyanobacterial origin are often not retargeted to the chloroplast but to various other compartments within the plant cell. Conversely, chloroplasts import proteins that were originally not encoded in the chloroplast genome (Martin et al., 2002).

After their lateral gene transfer into the nucleus, eukaryotic TCS genes have experienced remarkable innovations that are commonly not seen in bacterial TCS systems (Wuichet et al., 2010). First of all, the majority of eukaryotic TCSs belongs to the more complex MSP system, employing hybrid HKs and phosphotransfer proteins. The phosphotransfer proteins thereby act as phosphoryl shuttle to transmit the signal from the membrane into the nucleus. In order to cover such large distances without losing the signal, i. e. the phosphoryl group, phosphohistidines are employed instead of the much more labile phosphoaspartates. Secondly, TCSs can be integrated into other signaling cascades or regulatory networks, allowing for more complex regulation. For example, in case of the ethylene signaling TCS, the cognate ARRs act on the serine/threonine kinase CTR1, which then regulates transcription of respective target genes (Bisson and Groth, 2011), also by integration of MAP kinase cascades (Yoo et al., 2008). Finally, some TCS elements in plants lost their critical histidine or aspartic acid residues and can no longer function in a canonical His-Asp phosphorelay. For example, the red/far-red light sensing phytochromes or the CCT-family of pseudo-RRs involved in circadian clock function form such families of diverged HKs (Schaller et al., 2011). In the ethylene receptor family, canonical and diverged TCS elements co-exist and most likely also cooperate.

1.4 Specificity in plant TCS signaling

Recently, substrate specificity determining amino acid residues have been identified in the *E. coli* TCS, namely for the HK-RR pair of EnvZ-OmpR. First, a computational covariation approach revealed amino acids within the HK EnvZ that determine response regulator-specific targeting (Skerker et al., 2008). In a second step, trajectory-scanning mutagenesis of the EnvZ-OmpR interface helped to identify critical residues within the response regulator OmpR, which upon mutation changed recognition to non-cognate HKs (Capra et al., 2010). In a third study, the crystal structure of the phosphotransfer domain of $CheA_3$ ($CheA_3HP$) in complex with its cognate RR $CheY_6$ allowed for reengineering of the chemotaxis pathway in *R. sphaeroides* (Bell et al., 2010).

However, the results from the more basic TCS cannot be transferred onto the MSP system. In those more complex phosphorelays, the receiver domain is fused to the transmitter domain of the HK, thereby ensuring specific, *intra*molecular phosphoryl transfer. However, for the following *inter*molecular phosphoryl transfer from the HK onto the HP protein the situation is less clear. In case of *A. thaliana*, several independent studies indicate that HKs can phosphorylate multiple AHPs, and AHPs can receive phosphoryl groups from different HKs (Dortay et al., 2006; Hutchison et al., 2006). However, it is widely accepted that input specificity at the HK-level is maintained during phosphorelay signaling. Thus, although AHPs seem to serve as signaling hubs, there have to exist molecular mechanisms to ensure that signal output (of the ARRs) matches the signal input (from the HKs). This could include temporally and spatially separated expression patterns, as well as subtle differences in molecular recognition of single AHPs by upstream HKs and downstream ARRs. AHK4 dephosphorylates cognate AHPs *in vitro* and *in yeast* in a CK-dependent manner, and this therefore might also contribute to signal specificity in plants (Mahonen et al., 2006). However, it is unclear whether this phosphatase activity is physiologically relevant, as the used concentrations of AHKs and AHPs might not reflect the situation *in planta* (Kenney, 2010). Therefore, it is currently not known how signal specificity is maintained during plant MSP signaling.

Usually, TCSs are protected against cross-talk even from highly homologous TCS pathways (Groban et al., 2009). However, TCS pathways are not totally isolated but often encounter accessory proteins (Jung et al., 2012). These proteins are involved in co-sensing (Buelow and

Raivio, 2010; Herrou et al., 2010), physical assembly (Heermann et al., 2009), interconnection of TCS systems (Kox et al., 2000), and allosteric regulation (Paul et al., 2008). However, further studies are needed in order to reveal the complexity of TCS networks in a cellular context.

1.5 Significance of TCS signaling

1.5.1 TCS proteins as drug targets

The absence of TCS systems in animals and humans together with their established role in microbial pathogenicity makes bacterial TCS systems promising targets for drug development (Bourret and Silversmith, 2010). This is especially important as bacteria are becoming increasingly resistant to conventional antibiotics. Although of importance, antibiotic drug development is not restricted to clinical applications but might also become useful in agricultural crop farming to tackle plant pathogens. TCSs are particularly attractive drug targets as such drugs can be designed to act very selectively on specific TCS pathways or TCS proteins such as the sensor domains of HKs or a group of cognate RRs (Gotoh et al., 2010).

1.5.2 TCSs in synthetic biology

TCS elements are almost ubiquitous present in prokaryotic cells, and due to their functional modularity TCSs provide an immense repertoire of signaling modules for the construction of signaling networks (Ninfa, 2010). The design of novel connections between input stimuli and output responses can be exploited in the development of engineered cells (Ninfa, 2010). Such cells can be highly beneficial for medicinal and industrial applications as demonstrated for the synthesis of an anti-malaria drug precursor (Ro et al., 2006).

2 Introduction Part II: Adenoviruses

2.1 Classification and tropism

Adenoviruses (Ads) are widespread within the human population and usually cause mild infections of the gastrointestine, urinary tract, respiratory tract, or the eye (Arnberg, 2009). In immunocompromised patients, however, Ad infections are often lethal (Symeonidis et al., 2007). To date, over 50 human Ad serotypes are known and they are classified into seven species A-G (Table 1).

Table 1. Classification, tropism and primary receptors of Ad serotypes; adapted from Cupelli et al., 2011.

Species	Serotype	Infection	Primary receptor
A	12, 18, 31	Intestine	CAR
B	16, 21, 50	Respiratory tract, eye	CD46
	11, 34, 35	Respiratory and/or urinary tract, eye	CD46
	3, 7, 14	Respiratory and/or urinary tract, eye	DSG-2
C	1, 2, 5, 6	Respiratory tract	CAR
D	9, 10, 13, 15, 17, 19, 20, 22-25, 27-30, 32, 33, 36, 38, 39, 42-49, 51	Eye, intestine	CAR
	26		CD46
	8, 37, 64		Sialic acid, GD1a
E	4		CAR
F	40, 41	Intestine	CAR
G	52	Intestine	Sialic acid

Notably, some Ads do cause severe infections in otherwise healthy patients. In particular, Ad4, Ad8, Ad37, Ad53, Ad54 and Ad64 (formerly classified as Ad19a (Zhou et al., 2012)) are responsible for recurring outbreaks of epidemic keratoconjunctivitis (EKC) (Kaneko et al., 2011). This painful and highly contagious ocular disease results in decreased vision for several weeks and affects people all over the world (Akiyoshi et al., 2011; Fujimoto et al., 2012; Janicijevic-Petrovic et al., 2011; Mema et al., 2010; Meyer-Ruesenberg et al., 2011; Nercelles et al., 2010; Tabbara et al., 2010; Viney et al., 2008). As EKC spreads by contact, disease outbreaks are especially common in densely populated areas (e. g. in Asia) and in

medical wards with insufficient hygiene precautions (Kaltsas and Sepkowitz, 2012; Meyer-Ruesenberg et al., 2011). Although EKC is regarded as the most common viral eye disease worldwide (Meyer-Ruesenberg et al., 2011) with up to one million reported infections in Japan per year (Kaneko et al., 2011), currently no clinically approved antiviral treatment exists (Aplander et al., 2011).

2.2 Structure of the virion

Human Ads are large, non-enveloped viruses with a molecular weight of about 150 MDa and a diameter of approximately 90 nm. Ad virions have icosahedral symmetry and contain a 35-37 kbp, double-stranded DNA (dsDNA) genome. The viral particle shell, the capsid, is constructed from three major capsid proteins (hexon, penton base and fiber) and four minor capsid-associated proteins (IIIa, VI, VIII, and IX) (Russell, 2009) (Figure 6). Inside the capsid, the linear dsDNA genome is densely packed with three highly basic core proteins (V, VII, and μ) (Giberson et al., 2012), and this so-called adenosome also contributes to capsid stability (Smith et al., 2009). Further core proteins are the terminal protein TP, protein IVa2, and a viral protease. TP caps the 5'-DNA termini, whereas protein IVa2 is located at a single vertex and is involved in DNA packaging and capsid assembly (Christensen et al., 2008). The Ad protease is required for maturation of DNA-condensing proteins VII and μ, the core protein TP, and the minor structural proteins IIIa, VI, VIII (San Martin, 2012).
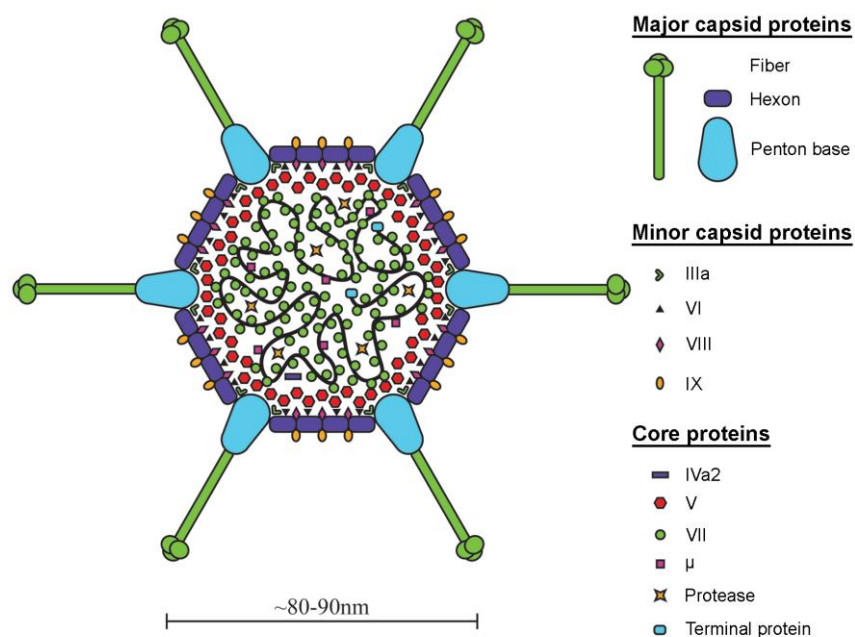

Figure 6. Organization of viral proteins in Ad virions; adapted from Giberson et al., 2012.

2.3 Organization of the capsid

The Ad capsid is primarily formed by 240 copies of the hexon protein, and the hexons are supplemented with one penton protein complex at each of the 12 vertices. The penton protein complex consists of the pentameric penton base protein and the trimeric fiber protein protruding from the fivefold axis of the penton base protein (Figure 7). Although the three major capsid proteins (hexon, penton base, and fiber) had been structurally characterized in isolation, the structural determinants for capsid assembly and stability remained elusive. Recently, Cryo-EM and X-ray structures of intact Ad virions allowed for the first time structural insight into the interactions between major and minor capsid proteins (Liu et al., 2010; Reddy et al., 2010). However, due to the still limited resolution of about 3.5 Å for both virion structures, not all visible secondary structure elements could be unambiguously assigned to one of the minor capsid proteins (Nemerow et al., 2012). For example, the crystal structure indicates that a four-helix bundle outside the capsid is part of protein IIIa, whereas the authors of the Cryo-EM structure assigned this four-helix bundle to the C-terminus of protein IX (Liu et al., 2010; Nemerow et al., 2012; Reddy et al., 2010).
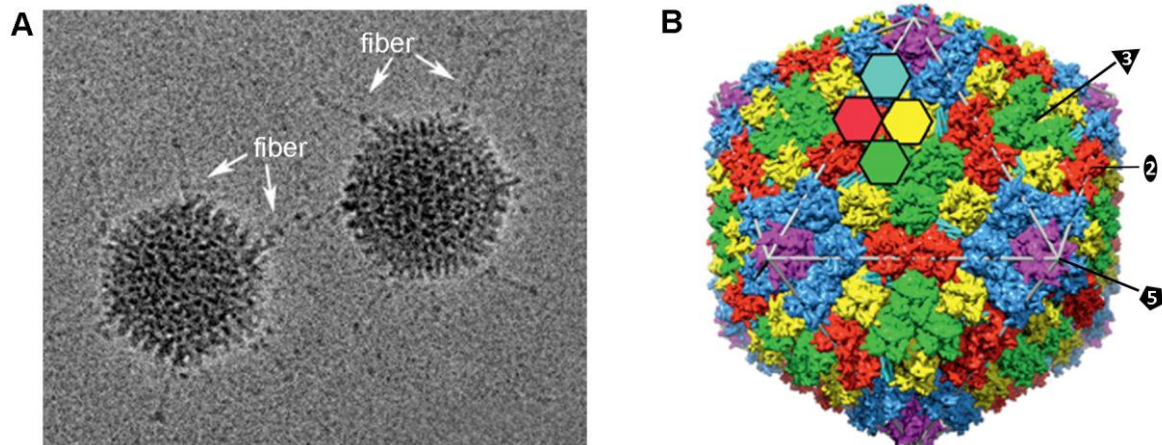

Figure 7. Overall structure and organization of the Ad capsid. (A) Cryo-electron micrograph of Ad5; adapted from Goosney et al., 2003. **(B)** Assembly of hexon and penton base proteins in the Ad capsid; adapted from Nemerow et al., 2012. The four unique hexon positions are represented by red, blue, green and yellow hexagons. The penton base protein is shown in purple. The 2-fold, 3-fold, and 5-fold symmetry axes are indicated. The fiber knob (not shown) protrudes at the vertices from the penton base protein along the 5-fold axes.

The minor capsid proteins fulfill diverse roles in capsid stabilization and at several stages during the adenoviral lifecycle. Protein IIIa is associated with stabilization of the vertices and plays a role in correct genome packaging (Ma and Hearing, 2011). Protein VI is a very

versatile protein involved in multiple processes throughout the Ad lifecycle. It plays roles in endosomal membrane lysis, viral trafficking inside the cell, adenoviral gene expression, transport of newly synthesized hexons, and it also interacts with the viral protease (San Martin, 2012). The remaining two proteins, VIII and IX, act as molecular glue between the hexon proteins (Liu et al., 2010). In addition, protein IX has also been associated with intracellular trafficking (Strunze et al., 2011)

2.3.1 The fiber protein

The trimeric fiber protein comprises a tail, a fibrous shaft, and a globular head domain called the knob. The tail anchors the fiber on the penton base protein, whereas the knob mediates primary cell attachment and can engage remarkably different cellular receptors (Cupelli and Stehle, 2011). The coxsackie- and adenovirus receptor (CAR) and the membrane cofactor protein CD46 attach to the side of the fiber knob, whereas sialic acid binds on top of the fiber knob. The binding site of desmoglein-2 (DSG-2) is still unknown. The shaft is built up by variable numbers of β-spiral repeats dependent on Ad serotype (Figure 8) (Chroboczek et al., 1995). Fiber lengths are ranging from about 120 Å to 340 Å, corresponding to 6 repeats in Ad3 and 22 repeats in Ad2 and Ad5, respectively (Ruigrok et al., 1990).
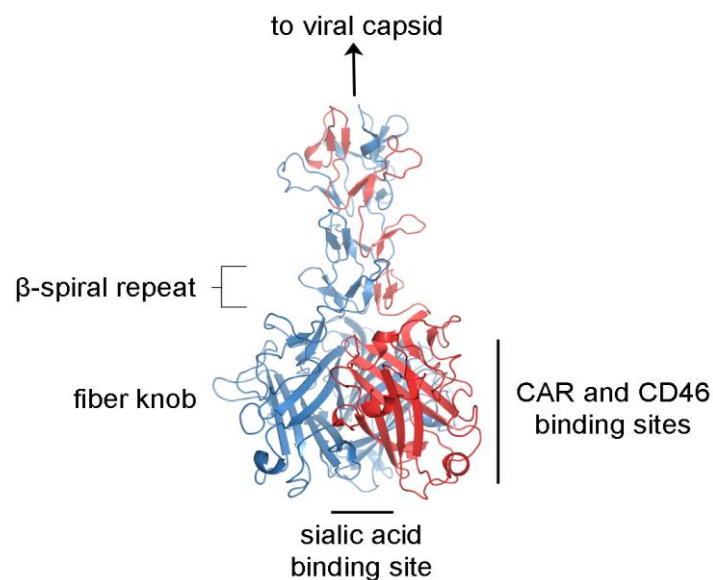

Figure 8. Crystal structure of the Ad2 fiber. The structure of the trimeric fiber is shown in cartoon representation and includes the knob and four β-spiral repeats of the shaft (van Raaij et al., 1999). One of the three protomers is highlighted in red. Regions for attachment of cellular receptors to Ads of all serotypes are indicated.

Interestingly, Cryo-EM analyses revealed that fiber shafts of Ad2, Ad5, and Ad12 contain a kink at a position corresponding to a four amino acid insertion in their third β-spiral repeat unit (Nemerow et al., 2009; Ruigrok et al., 1990). Although this atypical insertion is present in many species A, C, and E Ads, it is missing in certain species D Ads such as Ad37, resulting in straight and rigid fiber shafts (Chiu et al., 2001; Nemerow et al., 2009). Together with the large differences in fiber length, the kink is a key determinant of fiber flexibility. In a cellular context, fiber flexibility has also been linked to the capability of Ads to efficiently engage their respective cellular receptors for attachment and subsequent internalization (Chiu et al., 2001; Cupelli and Stehle, 2011). Interestingly, two Ad serotypes (Ad40 and Ad41) carry both long and short fibers (Kidd et al., 1993; Pieniazek et al., 1990) allowing for engagement of additional cellular receptors.

2.4 The adenoviral life cycle

The Ad lifecycle starts with attachment of the virus capsid to target cells, then proceeds with cell entry, intracellular trafficking to the nucleus, replication of the viral genome, and finally concludes in assembly and release of progeny virions.

For primary cellular attachment Ads use their protruding fiber proteins, allowing for attachment to various receptors dependent on Ad serotype. Cellular receptors include tight junction proteins, such as CAR (Bergelson et al., 1997) and the membrane glycoprotein DSG-2 (Wang et al., 2011), the apically expressed membrane cofactor protein CD46 (Marttila et al., 2005), and sialylated glycoproteins carrying the GD1a motif (Nilsson et al., 2011). After initial recruitment of the fiber proteins, cellular integrins engage the viral penton base protein. Some integrins (e. g. $\alpha_M\beta_2$, $\alpha_L\beta_2$, $\alpha_3\beta_1$) serve as secondary attachment receptors (Huang et al., 1996; Salone et al., 2003), whereas other integrins ($\alpha_v\beta_3$, $\alpha_v\beta_5$, $\alpha_5\beta_1$) trigger clathrin-mediated endocytosis (Li et al., 2001; Meier and Greber, 2003; Wickham et al., 1993), the primary internalization mechanism for Ad virions (Kobiler et al., 2012; Meier and Greber, 2003).

Following endocytosis, Ads escape from early endosomes by employing the viral protein VI that lyses the endosomal membrane (Wiethoff et al., 2005), a process triggered by acidification of the endosome (Marsh and Helenius, 2006). Next, the capsid is picked up by the cellular motor protein dynein, which carries the Ad capsid along the cytoskeleton to the microtubule-organizing center (MTOC) in close proximity to the nuclear pore complex (NPC)

(Bremner et al., 2009). The NPC first recruits the Ad particle from the MTOC under involvement of the nucleus export factor CRM1 (Kobiler et al., 2012). Next, the NPC mediates the final steps of DNA uncoating and then provides the gateway for nuclear import of viral DNA (Greber et al., 1997). Various proteins have been found to assist in viral DNA uptake, including nucleoporins, histone H1, importins (α, β, and 7), transportin and Hsp70 (Kobiler et al., 2012).

Replication of adenoviral DNA can be divided into an early and a late phase (Russell, 2000). Early genes are transcribed first in order to bypass the host immune system and to ensure optimal conditions for viral replication within the host cell, followed by viral DNA synthesis itself (Lenaerts et al., 2008). Next, late genes encoding for structural proteins are transcribed for viral DNA encapsidation and maturation (Lenaerts et al., 2008). After assembly, progeny virions are released during cell lysis induced by expression of the Ad death protein (Tollefson et al., 1996). A second escape mechanism exploits the production of excess fiber proteins. These extra fibers bind to CAR in the tight junctions, thereby disrupting the CAR-CAR homodimerization. Thus, the tight junctions open up and Ad virions can move across the epithelial cell barrier (Arnberg, 2009; Walters et al., 2002).

2.5 Significance of adenoviral research

Ad research is highly relevant for several reasons. First and most obviously, structural and functional knowledge of Ads allows for the identification of drug target sites for prevention or treatment of Ad-mediated diseases. This is especially important in case of immunocompromised patients and for treatment of EKC. Second, Ad vectors are currently the most frequently used vectors for human gene therapy and cancer therapy (Arnberg, 2009). However, major problems have been encountered during Ad vector therapy by the development of neutralizing antibodies and retargeting of the Ad vectors to the liver (Arnberg, 2009). Thus, for a regular and comprehensive application, more potent therapeutic vectors need to be developed. Third, as Ads employ the entire cell during their lifecycle, they are highly useful vehicles to obtain a detailed understanding of fundamental cellular functions. This has been highlighted by the discovery of RNA splicing (Berget et al., 1977; Ricciardi et al., 1981) and the research on molecular transport mechanisms (Radtke et al., 2006).

2.5.1 Antivirals

In order to treat Ad-caused diseases a number of different strategies have been employed, but until today no antiviral drug has been approved for treatment of EKC. Most tested compounds were often already in use against other viruses (Lenaerts et al., 2008; Skevaki et al., 2011), and therefore these drugs are often meant to target all Ads irrespective of serotype. However, some of these proposed anti-adenoviral drug candidates have only been tested in animal models so far. As the EKC-causing Ads do not replicate in animal models (the New Zealand rabbit model) or do not show clinical symptoms (the cotton rat model), the therapeutic relevance of these compounds for treatment of EKC remains unclear (Inoue et al., 2009; Kinchington et al., 2005).

For example, the two anti-HIV drugs Zalcitabine and Stavudine, which both inhibit the reverse transciptase of HIV, showed anti-adenoviral activity *in vitro* and in an animal model (Inoue et al., 2009; Uchio et al., 2007). These two compounds need to be tested in clinical studies. For topical treatment of EKC the combined application of the antiseptic povidone-iodine and the steroid dexamethasone has been suggested (Clement et al., 2011; Pelletier et al., 2009), but more comprehensive studies including placebo control patients are required for evaluation (Skevaki et al., 2011).

One approach for systemic treatment of EKC is based on acyclic nucleotide analogs targeting the viral DNA polymerase (Kinchington et al., 2005). Among these compounds, Cidofovir was thought to be the most promising candidate for several years (Figure 9A). Cidofovir had previously been licensed for treatment of human cytomegalovirus retinitis in AIDS patients (Safrin et al., 1997). Although Cidofovir efficiently cleared infection and inhibited spread of infection in initial clinical trials, it showed severe side effects in follow-up studies and was therefore not marketed (Kinchington et al., 2005). However, a lipid conjugate of Cidofovir with increased bioavailability, CMX001, did not show such limiting side effects in first clinical trials and is a promising candidate for a broad-spectrum antiviral drug against multiple dsDNA viruses including EKC-causing Ads (Jiang et al., 2010; Painter et al., 2012) (Figure 9B). In another study $NMSO_3$, a lipid-linked sulfated sialic acid, was shown to be more potent against all tested Ads than Cidofovir or Zalcitabine (Kaneko et al., 2001) (Figure 9C). Importantly, the tested Ads included the EKC-causing Ad serotypes 8 and 37. $NMSO_3$ is a non-toxic compound with antiviral activity against respiratory syncytial virus (RSV) (Kimura et al., 2000). It was proposed that $NMSO_3$ binds to the RSV-F glycoprotein and thereby

inhibits infection (Kimura et al., 2000). This could be similar to the situation in EKC-causing Ads. In 2004, the binding mode of α2,3-sialyllactose to EKC-causing Ad37 was determined by X-ray crystallography (Burmeister et al., 2004). One might speculate that $NMSO_3$ could be accommodated in a similar fashion as seen for α2,3-sialyllactose. However, a reorientation of the binding site in Ad37 would be required due to the larger sulfate group at position C_4 in $NMSO_3$, so that the exact binding mode remains elusive.
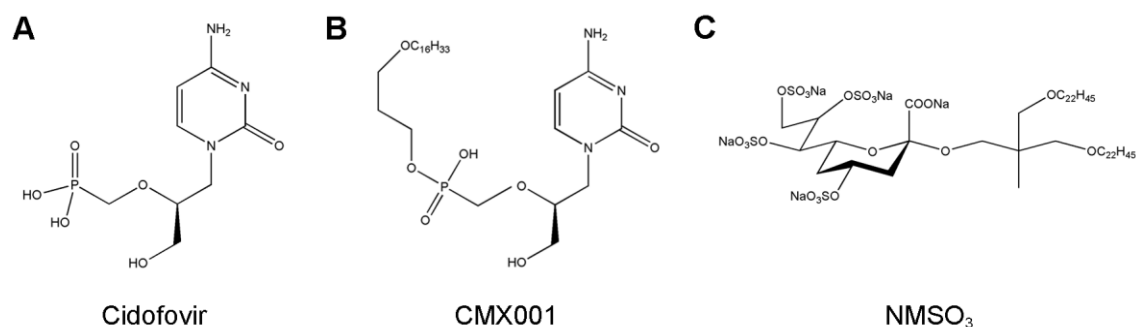

Figure 9. Chemical structure of anti-adenoviral drug candidates.

The discovery that sialic acid serves as a receptor for EKC-causing Ads led to the design of new, very specific strategies to combat EKC-causing Ads. In two studies, human serum albumin was functionalized with different numbers of linker molecules terminating in sialyllactose or sialic acid (Johansson et al., 2005; Johansson et al., 2007). The synthesized multivalent sialic acid conjugates prevented adenovirus type 37 from binding to human corneal epithelial, and also inhibited infection 1000-fold more efficiently than monovalent sialic acid (Johansson et al., 2007). However, as additional structure-guided modifications were introduced at the *N*-acetyl moiety of the sialic acid, this decreased the potency of the multivalent conjugates (Johansson et al., 2009).

Recently, the concept of aggregating Ad37 virions by multivalent sialic acid conjugates in order to inhibit infection was successfully extended. Thereby, sialic acid was linked via a polar linker to a lipid tail and then incorporated into liposomal sheets, which can mimic cellular membranes and are better suited sialic acid carriers from a pharmaceutical point of view (Aplander et al., 2011). These compounds still have to be evaluated in clinical trials.

3 Starting Point and Aims

3.1 MSP signaling in *A. thaliana*

Since the MSP system plays a central role in plant signaling, many MSP proteins have been functionally investigated in *A. thaliana* and other plants. However, when this project was started only two components of the Arabidopsis MSP system had been structurally characterized: the receiver domain of HK ETR1 (ETR1$_{RD}$) (Muller-Dieckmann et al., 1999) and the DNA-binding motif of ARR10 (Hosoda et al., 2002). In order to understand the plant MSP signaling system at the molecular level, the sensor HK AHK5 was chosen for structural and functional analysis.

At the beginning of this work, AHK5 was the least investigated HK in *A. thaliana* with no structural and only very little functional information being available. However, it had been shown that AHK5 is predominantly expressed in roots and weakly in flowers (Iwama et al., 2007). Additionally, AHK5 was known to mediate stomatal responses to exogenous and endogenous signals such as flagellin, which alter both the intracellular level of ROS and redox homeostasis (Desikan et al., 2008). The suggested role of AHK5 as an intracellular ROS sensor is in-line with the predicted lack of a transmembrane domain (Grefen and Harter, 2004). Accordingly, AHK5 is found in the cytoplasm but also co-localizes to the plasma membrane (Desikan et al., 2008). In order to determine the molecular basis of MSP signaling in plants we wanted to address the following questions:

- What is the structural basis of MSP signaling in plants?
- Which downstream AHPs does AHK5 recognize?
- What are the determinants of specificity for the AHK5-AHP interaction?
- What are the affinity and kinetics of the AHK5-AHP interaction?
- What are the implications for the mechanism of phosphoryl transfer?
- Does the MSP system use a conserved strategy for signaling?

In answering these questions, we sought to gain a detailed understanding of the molecular recognition events that underly MSP signaling in plants. Additionally, our results should define a useful platform for ongoing efforts to reengineer MSP signaling cascades.

3.2 Cellular attachment of Ad37

The identification of the GD1a glycan as a cellular receptor for Ad37 by glycan array screening and cell-culture based binding and infection studies provided the basis for the development of specific antivirals against EKC-causing Ads. Such antivirals would inhibit Ad37 binding to target cells, and thus would prevent spread of the infection. In order to develop such antivirals, we first sought to identify the binding determinants for the Ad37-GD1a interaction. This knowledge should then be exploited for structure-based drug design. Therefore, we wanted to address the following key questions:

- How does Ad37 engage the GD1a glycan?
- What are the determinants of specificity for the Ad37-GD1a interaction?
- How does the sialic acid binding mode in Ad37 differ from those seen in other viruses that use sialylated receptors?
- What opportunities does the Ad37-GD1a binding mode offer for structure-based drug design?
- How do those inhibitors bind to the Ad37 fiber knob?
- How can the potency of those Ad37 inhibitors be improved?

4   Results and Discussion

4.1 Structure-function analysis of *Arabidopsis thaliana* histidine kinase AHK5 bound to its cognate phosphotransfer protein AHP1

A key signal transduction pathway in plants and many eukaryotes proceeds via the MSP system. Within this MSP, external stimuli first lead to the activation of a HK, and this is then followed by transfer of a phosphoryl group from the receiver domain of the kinase (HK_{RD}) to downstream, cytosolic AHPs.

In order to establish the determinants of specificity for this signaling relay system in plants, the crystal structure of $AHK5_{RD}$ in complex with one of its cognate HPs, AHP1, was solved to 1.95 Å resolution. This is the first structure for an *A. thaliana* plant HP, and importantly the first structure for any plant MSP signaling complex. Inspection of the $AHK5_{RD}$-AHP1 interface revealed two prominent features, the hydrophobic patch and the hydrogen bond docking ridge. Although critical interface residues are conserved among all AHP proteins, the employed interaction strategy differs significantly from other structurally characterized prokaryotic and eukaryotic HPs.

In addition to the structural analysis, we have performed surface plasmon resonance (SPR) experiments, which for the first time provide a quantitative recording of plant MSP-signaling interactions in plants. These experiments show that $AHK5_{RD}$ engages different AHPs (AHP1, AHP2, and AHP3) with similar, micromolar affinity *in vitro*. In order to analyze this interaction also in the context of complete signaling cascades *in planta*, we conducted bimolecular fluorescence complementation (BiFC) assays. These semi-quantitative assays show differences in the $AHK5_{RD}$-AHP interaction strength, indicating the presence of (yet unknown) additional molecular mechanisms which affect and interfere with the $AHK5_{RD}$-AHP interaction *in planta*.

The correlation of structural and functional data provides the first insight, at the atomic level as well as with quantitative affinity data, into the molecular recognition events governing the MSP in plants. In combination with the established SPR protocol, the observed

interactions in the AHK5$_{RD}$-AHP1 complex structure moreover establishes a firm platform for ongoing efforts to reengineer and evaluate MSP signaling cascades.

4.2 The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis

Nilsson, E. C., Storm, R. J., **Bauer, J.**, Johansson, S. M. C., Lookene, A., Ångström, J., Hedenström, M., Eriksson, T. L., Frängsmyr, L., Rinaldi, S., Willison, H. J., Domelöff, F. P., Stehle, T., Arnberg, T. **(2011)**. The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. Nature Medicine, **17**, 105-109.

Ad37 is a leading cause of EKC, a severe and highly contagious ocular disease. Whereas most other Ads infect cells by engaging CD46, DSG-2 or CAR, Ad37 binds to sialic acid-containing cell surface molecules. However, the identity of the Ad37 receptor has remained unknown. Using glycan array screening, we found that the receptor-recognizing knob domain of the Ad37 fiber protein specifically binds to a branched hexasaccharide (the carbohydrate portion of the GD1a ganglioside) that features two terminal sialic acids. Soluble GD1a glycans and GD1a-binding antibodies efficiently prevented Ad37 virions from binding to and infecting human corneal epithelial (HCE) cells. Interestingly, further experiments showed that the GD1a glycan receptor motif rather belongs to multiple O-linked cell surface glycoproteins than to a glycolipid. X-ray crystallographic studies revealed that the two terminal sialic acid residues dock into two of three previously established sialic acid-binding sites in the trimeric Ad37 fiber knob. The detailed understanding of the structural parameters that guide recognition of the GD1a glycan by the Ad37 virion establishes the basis for the development of inhibitors that prevent cells from Ad37 attachment.

4.3 A potent trivalent sialic acid inhibitor of adenovirus type 37 infection of human corneal cells

Spjut S., Qian, W., **Bauer, J.**, Storm R., Frängsmyr, L., Stehle, T., Arnberg, N., Elofsson, M. **(2011)**. A Potent Trivalent Sialic Acid Inhibitor of Adenovirus Type 37 Infection of Human Corneal Cells. Angewandte Chemie Int. Ed. Engl., **50**, 6519-6521.

The asymmetry of binding in the Ad37-GD1a complex structure - a bivalent receptor versus three possible binding sites - was exploited for the design of a trivalent inhibitor (ME0322)

carrying three terminal sialic acid residues. This trivalent glycoconjugate efficiently inhibited binding of Ad37 to HCE cells as well as Ad37 infection. The co-crystal structure of the Ad37-inhibitor complex reveals that one inhibitor molecule engages now all three binding sites of the Ad37 fiber protein. Biological competition experiments demonstrate that ME0322 is a 1000-fold more potent inhibitor of Ad37 infection than soluble GD1a glycan, and also substantially more potent than 3'-sialyllactose-HSA and sialic acid-HSA conjugates. Thus, structure-function analysis of the virus-inhibitor interaction shows that such trivalent molecules may be suitable for topical treatment of EKC.

4.4 Viruses and sialic acids: rules of engagement

Neu, U., **Bauer, J.**, Stehle, T. **(2011)**. Viruses and sialic acids: rules of engagement. Current Opinion in Structural Biology, **21**, 610-618.

Proteins and lipids on eukaryotic cellular surfaces are largely decorated with glycans terminating in sialic acid residues (Varki, 2008). Such sialylated glycans serve in many cases as receptors for pathogenic viruses including influenza viruses, noroviruses, rotaviruses, polyomaviruses, and Ads (Erickson et al., 2009; Isa et al., 2006; Neu et al., 2009; Nilsson et al., 2011; Taube et al., 2009; Viswanathan et al., 2010). Within the last decade the pool of viral surface proteins complexed by sialylated glycan receptors has significantly expanded. This allowed us to compare the binding modes of sialylated receptors in representative structures of virus-receptor complexes in order to define common principles for sialic acid engagement. Importantly, viruses often bind to sialylated receptors primarily by contacting the terminal sialic acid residues. Moreover, virions generally engage sialic acid residues from the same side of the sialic acid ring by contacting the *N*-acetyl group and the carboxylate function. Other contacting epitopes of the sialic acid and also the linkage of the sialic acid to the neighbouring sugar residue in the oligosaccharide chain differ among viruses, thus ensuring specificity. Although sialic acid is rich in different epitopes *per se* compared to other monosaccharides, there exist over 40 additional sialic acid variants that allow for even more specific modes of recognition. In most cases, viruses engage a single sialic acid residue only with low affinity but increase the virus-host interaction strength by using multiple receptors. In case of Ad37, this strategy is modified by using a bivalent instead of a monovalent

receptor, thereby avoiding the need of receptor clustering on the cellular surface for productive infection.

4.5 Ongoing research

Our current research focuses on the improvement of the first trivalent Ad37 inhibitor ME0322 in order to increase the potency of such sialylated conjugates. Therefore, new inhibitors with a smaller scaffold were synthesized in the research group of Mikael Elofsson (Umeå, Sweden) (Figure 10).
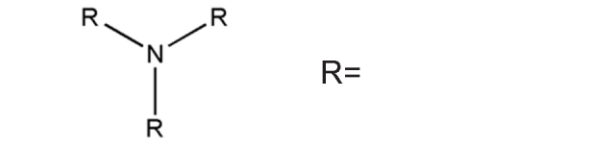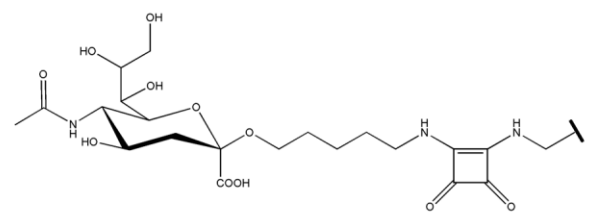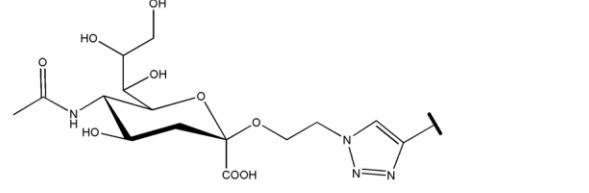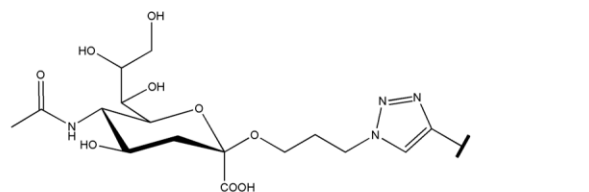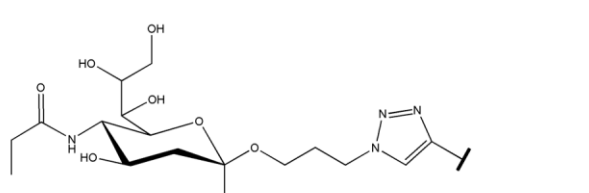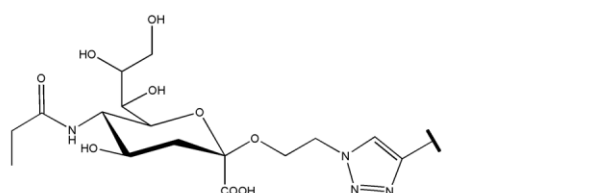

Figure 10. Overview of all trivalent Ad37 inhibitors examined here.

First, the squaric acid-scaffold in ME0322 was replaced by a triazole-scaffold, yielding conjugates ME0385 and ME0386. Both conjugates were then co-crystallized with the Ad37 fiber knob following the approach described previously (Spjut et al., 2011). In both complex structures, Ad37-ME0385 as well as Ad37-ME0386, electron density for the entire conjugate was visible, thus allowing for unambiguous placement of the ligands (Figure 11A-C and Table 2).

In a second round of inhibitor engineering, the *N*-acetyl groups in ME0385 and ME0386 were elongated to propanamide groups, resulting in conjugates ME0408 and ME0407, respectively. Ad37 complex structures were then obtained by soaking (using 2 mM ME0408) or co-crystallization (for ME0407) as described previously (Nilsson et al., 2011; Spjut et al., 2011). Overall, crystal structures for both conjugates show a highly similar binding mode compared to previous conjugates. The introduced propanamide moiety fills a hydrophobic pocket that is formed by Tyr308, Gly309, Tyr310, and Val322 of one protomer and is completed by Tyr312 and Pro317 of the neighbouring protomer (Figure 11D and Table 2). Due to the enlargement of the hydrophobic contact area, we expect an increase in affinity of the inhibitor. Although this strategy failed in case of similarly modified HSA-conjugates (Johansson et al., 2009), the situation might be different in the context of smaller conjugates with altered scaffolds. Currently, all four conjugates (ME0385, ME0386, ME0407, and ME0408) are evaluated in biological competition and binding experiments.
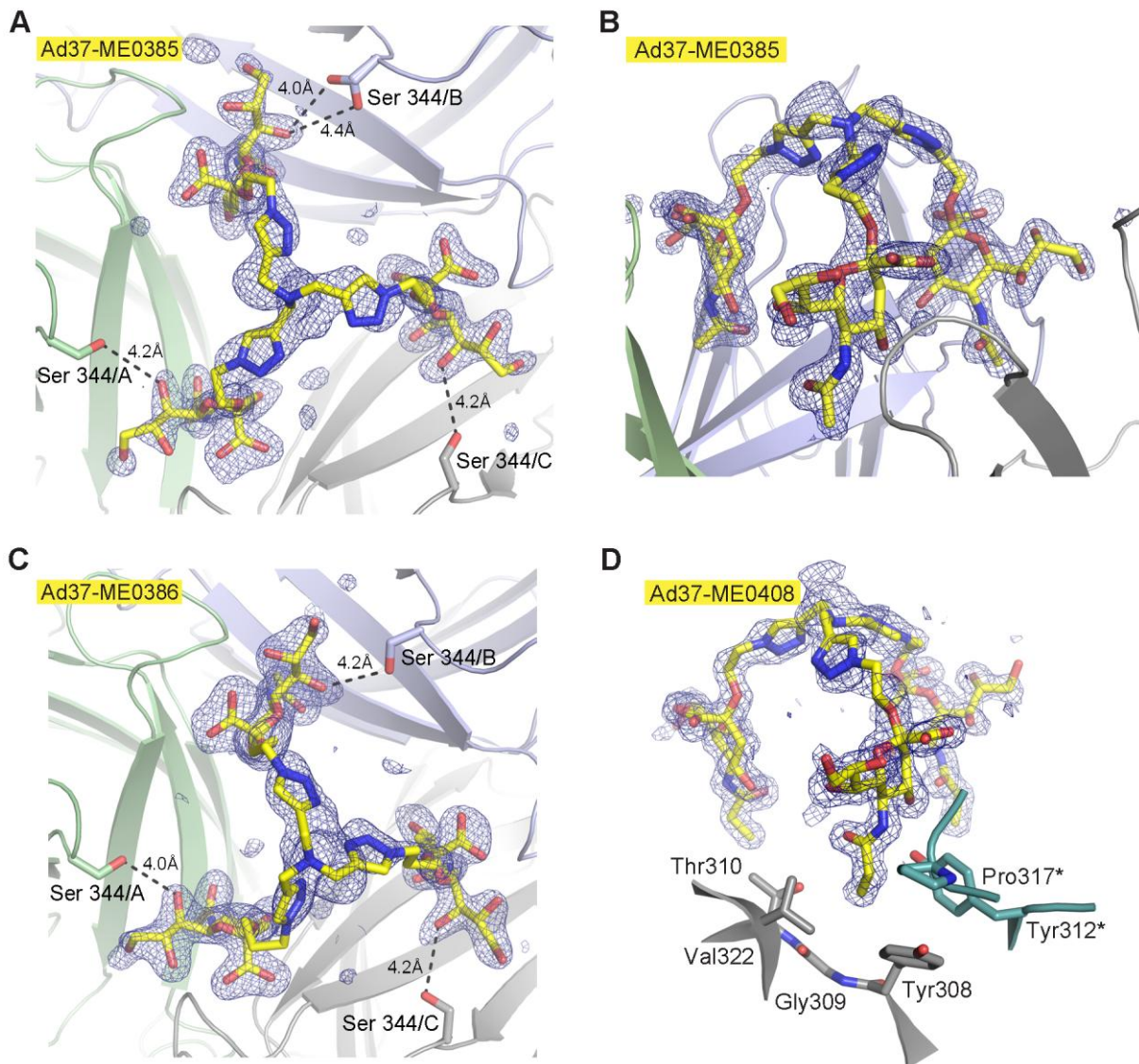
Figure 11. Binding of ME0385, ME0386, and ME0408 to the Ad37 fiber knob.

(A) and **(B)** Simulated annealing omit difference electron density maps for ME0385, calculated at 2.0 Å resolution, contoured at 3.0 σ and shown with a radius of 4 Å around ME0385. The knob is shown from top in (A) and from the side in (B). One ME0385 molecule (yellow) is bound with its three terminal sialic acids. Black dashed lines in (A) indicate distances to Ser344, which could be contacted in modified follow-up inhibitors.

(C) Simulated annealing omit difference electron density map for ME0386, calculated at 1.9 Å resolution, contoured at 3.0 σ and shown with a radius of 4 Å around ME0386. One ME0386 molecule (yellow) is bound with its three terminal sialic acids. Black dashed lines indicate distances to Ser344, which could be contacted in modified follow-up inhibitors.

(D) Simulated annealing omit difference electron density map for ME0408 (yellow), calculated at 1.5 Å resolution, contoured at 3.0 σ and shown with a radius of 4 Å around ME0408. One binding pocket for the propanamide moiety is shown; residues of one protomer are shown in grey stickes, residues of the neighbouring protomers are shown in cyan and marked with an asterisk.

Table 2. Data collection and refinement statistics.

	Ad37-ME0385	Ad37-ME0386	Ad37-ME0407	Ad37-ME0408
Data collection				
Beamline	X06SA (SLS)	X06SA (SLS)	X06DA (SLS)	X06DA (SLS)
Space group	P 2_1	P 2_1	P 2_1	P 2_1
Cell dimensions				
a, b, c (Å)	74.94, 67.18, 121.43	60.86, 69.91, 74.07	61.32, 69.68, 74.78	60.93, 69.28, 74.56
α, β, γ (°)	90.00, 96.81, 90.00	90.00, 94.62, 90.00	90.00, 94.43, 90.00	90.00, 94.92, 90.00
Ad trimer / ASU	2	1	1	1
Resolution (Å)	40-2.0 (2.05-2.00)	40-1.90 (1.95-1.90)	50-1.50 (1.54-1.50)	50-1.50 (1.54-1.50)
$CC_{1/2}$ (%)	98.6 (75.7)	99.6 (80.1)	99.9 (57.9)	99.9 (62.8)
$I / \sigma I$	6.1 (2.1)	9.9 (1.8)	14.6 (1.3)	15.1 (1.7)
Completeness (%)	99.3 (99.1)	96.8 (92.9)	99.5 (95.7)	99.4 (96.0)
Redundancy	3.0 (3.0)	2.9 (2.3)	3.7 (3.5)	3.7 (3.5)
Wilson B-factor (Å^2)	30.4	36.7	27.5	27.3
Refinement				
Resolution (Å)	37.26-2.00	37.91-1.90	45.96-1.50	45.66-1.50
No. of unique reflections	79104	47283	97974	96426
R_{work} / R_{free}	16.4 / 20.6	17.9 / 21.4	15.5 / 19.3	14.3 / 17.7
No. of non-H atoms				
Protein chain a / b / c	1466 / 1464 / 1462	1464 / 1463 / 1473	1473 / 1453 / 1473	1457 / 1461 / 1476
Protein chain e / f / g	1458 / 1451 / 1472	-	-	-
Inhibitor	88	91	94	91
Water	873	511	621	665
B-factors (Å^2)				
Protein chain a / b / c	20.8 / 21.7 / 25.8	40.8 / 37.3 / 29.2	25.4 / 26.8 / 19.0	26.0 / 26.7 / 18.4
Protein chain e / f / g	22.1 / 24.8 / 20.2	-	-	-
Inhibitor	34.2	45.9	24.9	21.2
Water	33.7	39.6	34.1	37.5
R.m.s. deviations				
Bond lengths (Å)	0.012	0.012	0.015	0.013
Bond angles (°)	1.333	1.349	1.446	1.444

* Values in parentheses are for highest resolution shell.

4.6 Outlook

4.6.1 Structure-function analysis of plant MSP systems

The structure-function analysis of $AHK5_{RD}$-AHP complexes defined the parameters that guide molecular recognition of $AHK5_{RD}$ by AHPs. Although interface residues in AHPs show a high degree of conservation, critical interface residues are less conserved among the plant HK_{RD} proteins. With the established SPR protocol at hand it would be interesting to evaluate whether and how differences in HK_{RD} interface residues translate into different HK_{RD}-AHP affinities. These *in vitro* experiments could be accompanied by comprehensive BiFC analysis *in planta*. Furthermore, the crystal structure of $AHK5_{RD}$-AHP1 represents the non-phosphorylated state of this signaling complex. By using other HK_{RD} proteins, it might be possible to successfully crystallize a complex in the phosphorylated state using phosphoryl mimics. Thus, one would be able to evaluate changes upon phosphorylation and confirm the proposed reaction mechanism.

In addition, our structure-function analysis revealed that recognition of the $AHK5_{RD}$ interface by AHPs does not impose specificity onto the MSP system alone. Thus, in order to maintain input specificity during signaling, other molecular mechanisms have to interfere with MSP signaling. This might include other regulatory proteins or differential temporary expression patterns of cognate AHPs and ARRs in response to a certain input stimulus. Determination of the cellular parameters that govern such regulation and specificity would provide valuable insight into MSP signaling. In addition, identified regulatory mechanisms might also serve as template for the identification of similar mechanisms in other signaling systems.

In order to gain a more comprehensive understanding of the complete MSP signaling cascade, the second intermolecular phosphoryl transfer step from the AHPs onto cognate ARRs could be analyzed. Structural and functional investigations of AHP-ARR interactions would allow for the determination of parameters that control AHP recognition by ARRs. Thus, one would be able to define differences in AHP recognition by HKs or ARRs, respectively. For investigation of these complexes one needs to screen for soluble ARR_{RD} proteins, which bind to one of the soluble AHP1-3 proteins. The key challenge will then consist in the crystallization of purified AHP-ARR_{RD} complexes.

4.6.2 Design of follow-up inhibitors against EKC-causing Ads

The structural analysis of Ad37 attachment to its cellular receptor GD1a led to the design of the trivalent inhibitor ME0322 (Figure 12). We then developed and structurally characterized further conjugates with smaller scaffolds and modified sialic acids. However, there are many more routes that can be taken in order to further improve the potency of such Ad37-inhibitors.
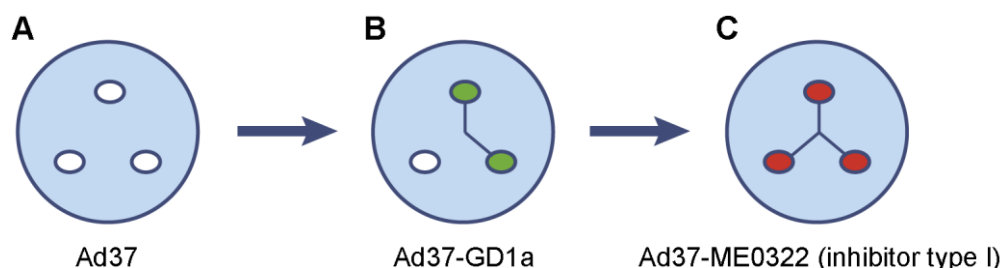

Figure 12. Scheme of Ad37 inhibitor development. (A) The Ad37 fiber knob has three free binding sites shown as white spheres. **(B)** The bivalent cellular receptor GD1a engages two of the three binding sites (shown as green spheres). **(C)** The inhibitor ME0322 engages all three Ad37 binding sites (red spheres) via a pyramid-like linker. ME0322 is a type I inhibitor and represents here all trivalent inhibitors examined in this work (ME0322, ME0385, ME0386, ME407, and ME408).

- Analysis of the electron density for all investigated conjugates revealed that the linker is flexible and can adopt multiple conformations. Therefore, the linker could be shortened by about 3 Å and thus become more rigid than it is in conjugates ME0385 or ME0408. Although a shorter linker would not help in contacting additional residues of the Ad37 knob, it might increase affinity by the higher local concentration of the sialic acid residues towards their respective binding sites.

- One could target the hydroxyl group of Ser344 by modification of the glycerol chain (Figure 11A, C). This strategy would result in three additional hydrogen bonds within the entire conjugate and is thus expected to significantly increase the affinity of such an inhibitor.

- In another approach, one could use a small circular scaffold, such as Kemp`s triacid (Figure 13A). Kemp`s triacid would be positioned on top of the central Ad37-cavity and could engage each of the three sialic acids by a short linker via functionalized carboxylate groups, resulting in a type II inhibitor (Figure 13B). Such a compound might also be beneficial as this might allow for a more rigid design of the linker, thus resulting in higher affinity binding.
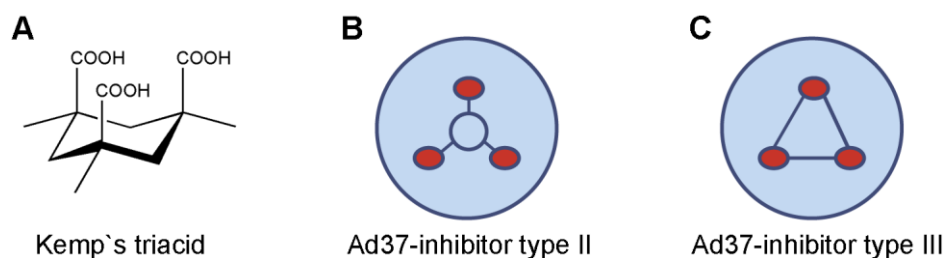
Figure 13. Circular scaffolds in Ad37-inhibitor design. (A) Chemical structure of Kemp`s triacid. **(B)** Scheme for type II inhibitors involving small circular scaffolds such as Kemp`s triacid. **(C)** Scheme for type III inhibitors that directly connect the sialic acid residues.

- The bridging scaffold in our series of Ad37-inhibitors points away from the fiber knob protein without any possibility to contact the Ad37 fiber knob from the bridging scaffold. Therefore, it would be highly beneficial to design an inhibitor that contacts the Ad37 knob protein and thus increases the affinity of the conjugate. Such a circular type III inhibitor could be formed by connecting one of the hydroxyl groups of the sialic acid glycerol chain with the carboxylate-function of the next sialic acid molecule (Figure 13C). However, this might result in decreased overall affinity as Lys345 makes a salt bridge with this carboxylate. Thus, it would be difficult to synthesize such a compound without compromising affinity. A more promising type III inhibitor could be designed by connecting the 4-OH group with the amide oxygen of the *N*-acetyl chain. The distance of about 8.5 Å would allow for a relatively short and rigid linker. In addition, this linker could be designed in a way that three additional hydrogen bonds are formed to the side chains of Thr310 in the Ad37 trimer.

- Very recently, several monovalent sialic acid conjugates were incorporated into liposomal sheets in order to aggregate Ad37 virions (Aplander et al., 2011). By exchanging the monovalent sialic acids by the more potent trivalent inhibitors, Ad37 virions could be even more efficiently captured. Since trivalent conjugates (IC_{50} for ME0322 = 0.4 µM) were four orders of magnitude more potent than monovalent sialic acid in inhibiting Ad37 infection (Spjut et al., 2011), a similar decrease for trivalent conjugates incorporated into liposomal sheets might result in an IC_{50}-value of about 70 pM. Such a compound with medium picomolar affinity would mean an immense increase in inhibitory power over those of existing drug candidates for antiviral treatment of EKC.

5   Summary

This thesis investigates the structural and functional basis of the plant two-component system and the cellular attachment of Ad37. In case of the plant two-component system, the crystal structure of AHK5$_{RD}$-AHP1 provided for the first time insight into the binding and phosphoryl transfer mechanism of the sensor HK AHK5 to its downstream signaling partner, AHP1. Analysis of the binding interface in combination with sequence comparison of AHP1 to the remaining AHP proteins in *A. thaliana* suggested that AHK5 would be able to engage any of the AHP proteins in a similar fashion. SPR analyses confirmed that AHK5$_{RD}$ indeed binds AHP1-3 with similar, micromolar affinity *in vitro*. *In planta*, however, BiFC studies revealed differences in the AHK5$_{RD}$-AHP interaction strength. This indicates the presence of (yet unknown) additional molecular mechanisms that modulate the AHK5$_{RD}$-AHP interaction *in planta*. Together with the established SPR protocol the structural analysis of the AHK5$_{RD}$-AHP1 complex provides a firm basis for further characterization of signaling in plant TCSs.

The discovery of the GD1a glycan as cellular receptor for Ad37 led to the determination of the crystal structure of the Ad37-GD1a complex. In this structure, the branched GD1a glycan docks with its two terminal sialic acid residues into two of three binding sites in a central cavity on top of the Ad37 fiber knob. This symmetry mismatch in binding, i. e. having three possible binding sites but only a bivalent receptor, was exploited to design the trivalent inhibitor ME0322. The ME0322 compound carries three branches, each decorated with a terminal sialic acid residue, and it thus engages all three binding sites simultaneously in the Ad37-ME0322 crystal structure. Binding, infection and biological competition experiments then demonstrated the potency of ME0322. In subsequent cycles of structure-based inhibitor design, the scaffold of the conjugate and the sialic acids were modified in order to increase the affinity of follow-up inhibitors. Taken together, structure-based drug design of trivalent inhibitors serves as excellent platform for the development of potent antiviral drugs to combat Ad37 and other EKC-causing Ads.

6 Zusammenfassung

Diese Doktorarbeit untersucht die strukturelle und funktionelle Grundlage des pflanzlichen Zweikomponenten-Systems und der zellulären Bindung von Ad37. Im Fall des pflanzlichen Zweikomponenten-Systems gewährte die Kristallstruktur von AHK5$_{RD}$-AHP1 zum ersten Mal Einblick in den Mechanismus der Bindung und der Phosphorylgruppen-Übertragung der Sensor HK AHK5 zu ihrem nachgeschalteten Partner in der Signalübertragung, AHP1. Die Analyse der Interaktionsfläche in Kombination mit Sequenzvergleichen von AHP1 zu den übrigen AHP Proteinen in *A. thaliana* legte nahe, dass AHK5 in der Lage sei jedes der AHP Proteine auf ähnliche Weise zu binden. SPR-Analysen bestätigten, dass AHK5 tatsächlich AHP1-3 *in vitro* mit ähnlicher, micromolarer Affinität bindet. BiFC-Studien *in planta* zeigten jedoch Unterschiede in der Interaktionsstärke der AHK5$_{RD}$-AHP Komplexe. Die deutet auf die Anwesenheit weiterer (noch unbekannter) molekularer Mechanismen hin, die die AHK5$_{RD}$-AHP1 Interaktion *in planta* beeinflussen. Zusammen mit dem etablierten Protokoll für SPR-Messungen bietet die Strukuranalyse des AHK5$_{RD}$-AHP1 Komplexes eine verlässliche Grundlage für eine weitergehende Charakterisierung der Signalübertragung in pflanzlichen Zweikomponenten-Systemen.

Die Entdeckung von GD1a als zellulärem Rezeptor für Ad37 führte zur Bestimmung der Kristallstruktur des Ad37-GD1a Komplexes. In dieser Struktur bindet das verzweigte GD1a Glykan mit seinen zwei terminalen Sialinsäure-Resten in zwei der drei Bindestellen in einer zentralen Einstülpung oben auf dem Ad37 Fiber. Diese symmetrische Diskrepanz im Bindeverhalten, d. h. das Vorliegen von drei Bindestellen aber einem nur zweiwertigen Rezeptor, wurde ausgenutzt den dreiwertigen Inhibitor ME0322 zu entwerfen. ME0322 besitzt drei Arme, jeder ist mit einer terminalen Sialinsäure bestückt, und es besetzt daher in der Ad37-ME0322 Kristallstruktur alle drei Bindestellen auf einmal. Binde -, Infektions -, und biologische Kompetitions-Experimente zeigten die Wirksamkeit von ME0322. In darauffolgenden Zyklen strukturbasierten Inhibitor-Designs wurde das Gerüst des Konjugates sowie die Sialinsäuren modifiziert um die Affinität nachfolgender Inhibitoren zu erhöhen. Zusammenfassend lässt sich sagen, dass strukturbasiertes Wirkstoffdesign dreiwertiger Inhibitoren eine exzellente Plattform für die Entwicklung wirksamer, antiviraler Medikamente darstellt, um Ad37 und andere EKC-auslösende Adenoviren zu bekämpfen.

7 References

Abdallah, F., Salamini, F., and Leister, D. (2000). A prediction of the size and evolutionary origin of the proteome of chloroplasts of Arabidopsis. Trends in Plant Science *5*, 141-142.

Abeles, F.B., Morgan, P.W., and Saltveit, M.E., Jr. (1992). Ethylene in plant biology, Second edition.

Akiyoshi, K., Suga, T., Fukui, K., Taniguchi, K., Okabe, N., and Fujimoto, T. (2011). Outbreak of Epidemic Keratoconjunctivitis Caused by Adenovirus Type 54 in a Nursery School in Kobe City, Japan in 2008. Japanese Journal of Infectious Diseases *64*, 353-355.

Alm, E., Huang, K., and Arkin, A. (2006). The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. Plos Computational Biology *2*, 1329-1342.

Aplander, K., Marttila, M., Manner, S., Arnberg, N., Sterner, O., and Ellervik, U. (2011). Molecular Wipes: Application to Epidemic Keratoconjuctivitis. Journal of Medicinal Chemistry *54*, 6670-6675.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature *408*, 796-815.

Argueso, C.T., Raines, T., and Kieber, J.J. (2010). Cytokinin signaling and transcriptional networks. Current Opinion in Plant Biology *13*, 533-539.

Arnberg, N. (2009). Adenovirus receptors: implications for tropism, treatment and targeting. Reviews in Medical Virology *19*, 165-178.

Bauer, J., Reiss, K., Veerabagu, M., Heunemann, M., Harter, K., and Stehle, T. (2013). Structure-function analysis of Arabidopsis thaliana histidine kinase AHK5 bound to its cognate phosphotransfer protein AHP1. Molecular plant, doi: 10.1093/mp/sss126.

Bell, C.H., Porter, S.L., Strawson, A., Stuart, D.I., and Armitage, J.P. (2010). Using Structural Information to Change the Phosphotransfer Specificity of a Two-Component Chemotaxis Signalling Complex. Plos Biology, e1000306.

Bergelson, J.M., Cunningham, J.A., Droguett, G., KurtJones, E.A., Krithivas, A., Hong, J.S., Horwitz, M.S., Crowell, R.L., and Finberg, R.W. (1997). Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5. Science *275*, 1320-1323.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A *74*, 3171-3175.

Binder, B.M., Rodriguez, F.I., and Bleecker, A.B. (2010). The Copper Transporter RAN1 Is Essential for Biogenesis of Ethylene Receptors in Arabidopsis. Journal of Biological Chemistry *285*, 37263-37270.

Bisson, M.M.A., Bleckmann, A., Allekotte, S., and Groth, G. (2009). EIN2, the central regulator of ethylene signalling, is localized at the ER membrane where it interacts with the ethylene receptor ETR1. Biochemical Journal *424*, 1-6.

Bisson, M.M.A., and Groth, G. (2011). New paradigm in ethylene signaling: EIN2, the central regulator of the signaling pathway, interacts directly with the upstream receptors. Plant signaling & behavior *6*, 164-166.

Bourret, R.B., and Silversmith, R.E. (2010). Two-component signal transduction. Current Opinion in Microbiology *13*, 113-115.

Bremner, K.H., Scherer, J., Yi, J., Vershinin, M., Gross, S.P., and Vallee, R.B. (2009). Adenovirus Transport via Direct Interaction of Cytoplasmic Dynein with the Viral Capsid Hexon Subunit. Cell Host & Microbe *6*, 523-535.

Buelow, D.R., and Raivio, T.L. (2010). Three (and more) component regulatory systems - auxiliary regulators of bacterial histidine kinases. Molecular Microbiology *75*, 547-566.

Burmeister, W.P., Guilligay, D., Cusack, S., Wadell, G., and Arnberg, N. (2004). Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. Journal of Virology *78*, 7727-7736.

Caesar, K., Thamm, A.M.K., Witthoeft, J., Elgass, K., Huppenberger, P., Grefen, C., Horak, J., and Harter, K. (2011). Evidence for the localization of the Arabidopsis cytokinin receptors AHK3 and AHK4 in the endoplasmic reticulum. Journal of Experimental Botany *62*, 5571-5580.

Capra, E.J., and Laub, M.T. (2012). Evolution of two-component signal transduction systems. Annual review of microbiology *66*, 325-347.

Capra, E.J., Perchuk, B.S., Lubin, E.A., Ashenberg, O., Skerker, J.M., and Laub, M.T. (2010). Systematic Dissection and Trajectory-Scanning Mutagenesis of the Molecular Interface That Ensures Specificity of Two-Component Signaling Pathways. Plos Genetics, e1001220.

Chen, N., Liu, Y., Liu, X., Chai, J., Hu, Z., Guo, G., and Liu, H. (2009). Enhanced Tolerance to Water Deficit and Salinity Stress in Transgenic Lycium barbarum L. Plants Ectopically Expressing ATHK1, an Arabidopsis thaliana Histidine Kinase Gene. Plant Molecular Biology Reporter *27*, 321-333.

Chen, Y.-F., Gao, Z., Kerris, R.J., III, Wang, W., Binder, B.M., and Schaller, G.E. (2010). Ethylene Receptors Function as Components of High-Molecular-Mass Protein Complexes in Arabidopsis. Plos One, e8640.

Chen, Y.F., Randlett, M.D., Findell, J.L., and Schaller, G.E. (2002). Localization of the ethylene receptor ETR1 to the endoplasmic reticulum of Arabidopsis. Journal of Biological Chemistry *277*, 19861-19866.

Chiu, C.Y., Wu, E., Brown, S.L., Von Seggern, D.J., Nemerow, G.R., and Stewart, P.L. (2001). Structural analysis of a fiber-pseudotyped adenovirus with ocular tropism suggests differential modes of cell receptor interactions. Journal of Virology *75*, 5375-5380.

Christensen, J.B., Byrd, S.A., Walker, A.K., Strahler, J.R., Andrews, P.C., and Imperiale, M.J. (2008). Presence of the adenovirus IVa2 protein at a single vertex of the mature virion. Journal of Virology *82*, 9086-9093.

Chroboczek, J., Ruigrok, R.W., and Cusack, S. (1995). Adenovirus fiber. Current topics in microbiology and immunology *199 (Pt 1)*, 163-200.

Clement, C., Capriotti, J.A., Kumar, M., Hobden, J.A., Foster, T.P., Bhattacharjee, P.S., Thompson, H.W., Mahmud, R., Liang, B., and Hill, J.M. (2011). Clinical and Antiviral Efficacy of an Ophthalmic Formulation of Dexamethasone Povidone-Iodine in a Rabbit Model of Adenoviral Keratoconjunctivitis. Invest Ophthalmol Vis Sci *52*, 339-344.

Cupelli, K., and Stehle, T. (2011). Viral attachment strategies: the many faces of adenoviruses. Current opinion in virology *1*, 84-91.

Deng, Y., Dong, H., Mu, J., Ren, B., Zheng, B., Ji, Z., Yang, W.-C., Liang, Y., and Zuo, J. (2010). Arabidopsis Histidine Kinase CKI1 Acts Upstream of Histidine Phosphotransfer Proteins to Regulate Female Gametophyte Development and Vegetative Growth. Plant Cell *22*, 1232-1248.

Desikan, R., Horak, J., Chaban, C., Mira-Rodado, V., Witthoeft, J., Elgass, K., Grefen, C., Cheung, M.-K., Meixner, A.J., Hooley, R.*, et al.* (2008). The Histidine Kinase AHK5 Integrates Endogenous and Environmental Signals in Arabidopsis Guard Cells. Plos One, e2491.

Dong, C.-H., Jang, M., Scharein, B., Malach, A., Rivarola, M., Liesch, J., Groth, G., Hwang, I., and Chang, C. (2010). Molecular Association of the Arabidopsis ETR1 Ethylene Receptor and a Regulator of Ethylene Signaling, RTE1. Journal of Biological Chemistry *285*, 40706-40713.

Dong, C.-H., Rivarola, M., Resnick, J.S., Maggin, B.D., and Chang, C. (2008). Subcellular co-localization of Arabidopsis RTE1 and ETR1 supports a regulatory role for RTE1 in ETR1 ethylene signaling. Plant Journal *53*, 275-286.

Dortay, H., Mehnert, N., Buerkle, L., Schmuelling, T., and Heyl, A. (2006). Analysis of protein interactions within the cytokinin-signaling pathway of Arabidopsis thaliana. FEBS Journal *273*, 4631-4644.

Erickson, K.D., Garcea, R.L., and Tsai, B. (2009). Ganglioside GT1b Is a Putative Host Cell Receptor for the Merkel Cell Polyomavirus. Journal of Virology *83*, 10275-10279.

Fujimoto, T., Matsushima, Y., Shimizu, H., Ishimaru, Y., Kano, A., Nakajima, E., Adhikary, A.K., Hanaoka, N., and Okabe, N. (2012). A Molecular Epidemiologic Study of Human Adenovirus Type 8 Isolates Causing Epidemic Keratoconjunctivitis in Kawasaki City, Japan in 2011. Japanese Journal of Infectious Diseases *65*, 260-263.

Gamble, R.L., Coonfield, M.L., and Schaller, G.E. (1998). Histidine kinase activity of the ETR1 ethylene receptor from Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America *95*, 7825-7829.

Gao, Z., Wen, C.-K., Binder, B.M., Chen, Y.-F., Chang, J., Chiang, Y.-H., Ill, R.J.K., Chang, C., and Schaller, G.E. (2008). Heteromeric interactions among ethylene receptors mediate signaling in Arabidopsis. Journal of Biological Chemistry *283*, 23801-23810.

Gattolin, S., Alandete-Saez, M., Elliott, K., Gonzalez-Carranza, Z., Naomab, E., Powell, C., and Roberts, J.A. (2006). Spatial and temporal expression of the response regulators ARR22 and ARR24 in Arabidopsis thaliana. Journal of Experimental Botany *57*, 4225-4233.

Giberson, A.N., Davidson, A.R., and Parks, R.J. (2012). Chromatin structure of adenovirus DNA throughout infection. Nucleic Acids Research *40*, 2369-2376.

Gotoh, Y., Eguchi, Y., Watanabe, T., Okamoto, S., Doi, A., and Utsumi, R. (2010). Two-component signal transduction as potential drug targets in pathogenic bacteria. Current Opinion in Microbiology *13*, 232-239.

Greber, U.F., Suomalainen, M., Stidwill, R.P., Boucke, K., Ebersold, M.W., and Helenius, A. (1997). The role of the nuclear pore complex in adenovirus DNA entry. Embo Journal *16*, 5998-6007.

Grefen, C., and Harter, K. (2004). Plant two-component systems: principles, functions, complexity and cross talk. Planta *219*, 733-742.

Grefen, C., Staedele, K., Ruzicka, K., Obrdlik, P., Harter, K., and Horak, J. (2008). Subcellular localization and in vivo interactions of the Arabidopsis thaliana ethylene receptor family members. Molecular plant *1*, 308-320.

Groban, E.S., Clarke, E.J., Salis, H.M., Miller, S.M., and Voigt, C.A. (2009). Kinetic Buffering of Cross Talk between Bacterial Two-Component Sensors. Journal of Molecular Biology *390*, 380-393.

Heermann, R., Weber, A., Mayer, B., Ott, M., Hauser, E., Gabriel, G., Pirch, T., and Jung, K. (2009). The Universal Stress Protein UspC Scaffolds the KdpD/KdpE Signaling Cascade of Escherichia coli under Salt Stress. Journal of Molecular Biology *386*, 134-148.

Hejatko, J., Pernisova, M., Eneva, T., Palme, K., and Brzobohaty, B. (2003). The putative sensor histidine kinase CKI1 is involved in female gametophyte development in Arabidopsis. Molecular Genetics and Genomics *269*, 443-453.

Herrou, J., Foreman, R., Fiebig, A., and Crosson, S. (2010). A structural model of anti-anti-sigma inhibition by a two-component receiver domain: the PhyR stress response regulator. Molecular Microbiology *78*, 290-304.

Horak, J., Janda, L., Pekarova, B., and Hejatko, J. (2011). Molecular Mechanisms of Signalling Specificity Via Phosphorelay Pathways in Arabidopsis. Current Protein & Peptide Science *12*, 126-136.

Hosoda, K., Imamura, A., Katoh, E., Hatta, T., Tachiki, M., Yamada, H., Mizuno, T., and Yamazaki, T. (2002). Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators. Plant Cell *14*, 2015-2029.

Hothorn, M., Dabi, T., and Chory, J. (2011). Structural basis for cytokinin recognition by Arabidopsis thaliana histidine kinase 4. Nature Chemical Biology *7*, 766-768.

Hua, J., and Meyerowitz, E.M. (1998). Ethylene responses are negatively regulated by a receptor gene family in Arabidopsis thaliana. Cell *94*, 261-271.

Huang, S., Kamata, T., Takada, Y., Ruggeri, Z.M., and Nemerow, G.R. (1996). Adenovirus interaction with distinct integrins mediates separate events in cell entry and gene delivery to hematopoietic cells. Journal of Virology *70*, 4502-4508.

Hutchison, C.E., Li, J., Argueso, C., Gonzalez, M., Lee, E., Lewis, M.W., Maxwell, B.B., Perdue, T.D., Schaller, G.E., Alonso, J.M*., et al.* (2006). The Arabidopsis histidine phosphotransfer proteins are redundant positive regulators of cytokinin signaling. Plant Cell *18*, 3073-3087.

Hwang, I., Sheen, J., and Muller, B. (2012). Cytokinin signaling networks. Annual review of plant biology *63*, 353-380.

Inoue, H., Sonoda, K.-h., Ishikawa, M., Kadonosono, K., and Uchio, E. (2009). Clinical Evaluation of Local Ocular Toxicity in Candidate Anti-Adenoviral Agents in vivo. Ophthalmologica *223*, 233-238.

Isa, P., Arias, C.F., and Lopez, S. (2006). Role of sialic acids in rotavirus infection. Glycoconjugate Journal *23*, 27-37.

Iwama, A., Yamashino, T., Tanaka, Y., Sakakibara, H., Kakimoto, T., Sato, S., Kato, T., Tabata, S., Nagatani, A., and Mizuno, T. (2007). AHK5 histidine kinase regulates root elongation through an ETR1-dependent abscisic acid and ethylene signaling pathway in Arabidopsis thaliana. Plant and Cell Physiology *48*, 375-380.

Janicijevic-Petrovic, M.A., Sreckovic, S., Petrovic, N., and Sarenac, T. (2011). Epidemic Keratoconjunctivitis. Srpski Arhiv Za Celokupno Lekarstvo *139*, 282-285.

Jiang, Z.-G., Cohen, J., Marshall, L.J., and Major, E.O. (2010). Hexadecyloxypropyl-Cidofovir (CMX001) Suppresses JC Virus Replication in Human Fetal Brain SVG Cell Cultures. Antimicrobial Agents and Chemotherapy *54*, 4723-4732.

Johansson, S., Nilsson, E., Qian, W., Guilligay, D., Crepin, T., Cusack, S., Arnberg, N., and Elofsson, M. (2009). Design, Synthesis, and Evaluation of N-Acyl Modified Sialic Acids as Inhibitors of Adenoviruses Causing Epidemic Keratoconjunctivitis. Journal of Medicinal Chemistry *52*, 3666-3678.

Johansson, S.M.C., Arnberg, N., Elofsson, M., Wadell, G., and Kihlberg, J. (2005). Multivalent HSA conjugates of 3 '-siallyllactose are potent inhibitors of adenoviral cell attachment and infection. Chembiochem *6*, 358-364.

Johansson, S.M.C., Nilsson, E.C., Elofsson, M., Ahlskog, N., Kihlberg, J., and Arnberg, N. (2007). Multivalent sialic acid conjugates inhibit adenovirus type 37 from binding to and infecting human corneal epithelial cells. Antiviral Research *73*, 92-100.

Jung, K., Fried, L., Behr, S., and Heermann, R. (2012). Histidine kinases and response regulators in networks. Current Opinion in Microbiology *15*, 118-124.

Kaltsas, A., and Sepkowitz, K. (2012). Community acquired respiratory and gastrointestinal viral infections: challenges in the immunocompromised host. Current Opinion in Infectious Diseases *25*, 423-430.

Kaneko, H., Kato, K., Mori, S., and Shigeta, S. (2001). Antiviral activity of NMSO3 against adenovirus in vitro. Antiviral Research *52*, 281-288.

Kaneko, H., Suzutani, T., Aoki, K., Kitaichi, N., Ishida, S., Ishiko, H., Ohashi, T., Okamoto, S., Nakagawa, H., Hinokuma, R*., et al.* (2011). Epidemiological and virological features of epidemic keratoconjunctivitis due to new human adenovirus type 54 in Japan. British Journal of Ophthalmology *95*, 32-36.

Kenney, L.J. (2010). How important is the phosphatase activity of sensor kinases? Current Opinion in Microbiology *13*, 168-176.

Kidd, A.H., Chroboczek, J., Cusack, S., and Ruigrok, R.W.H. (1993). Adenovirus type-40 virions contain 2 distinct fibers. Virology *192*, 73-84.

Kieber, J.J., Rothenberg, M., Roman, G., Feldmann, K.A., and Ecker, J.R. (1993). CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the Raf family of protein-kinases Cell *72*, 427-441.

Kimura, K., Mori, S., Tomita, K., Ohno, K., Takahashi, K., Shigeta, S., and Terada, M. (2000). Antiviral activity of NMSO3 against respiratory syncytial virus infection in vitro and in vivo. Antiviral Research *47*, 41-51.

Kinchington, P.R., Romanowski, E.G., and Jerold Gordon, Y. (2005). Prospects for adenovirus antivirals. Journal of Antimicrobial Chemotherapy *55*, 424-429.

Kobiler, O., Drayman, N., Butin-Israeli, V., and Oppenheim, A. (2012). Virus strategies for passing the nuclear envelope barrier. Nucleus (Austin, Tex) *3*, 526-539.

Kox, L.F.F., Wosten, M., and Groisman, E.A. (2000). A small protein that mediates the activation of a two-component system by another two-component system. Embo Journal *19*, 1861-1872.

Laub, M.T., and Goulian, M. (2007). Specificity in two-component signal transduction pathways. In Annual Review of Genetics (Palo Alto: Annual Reviews), pp. 121-145.

Lenaerts, L., De Clercq, E., and Naesens, L. (2008). Clinical features and treatment of adenovirus infections. Reviews in Medical Virology *18*, 357-374.

Li, E., Brown, S.L., Stupack, D.G., Puente, X.S., Cheresh, D.A., and Nemerow, G.R. (2001). Integrin alpha nu beta 1 is an adenovirus coreceptor. Journal of Virology *75*, 5405-5409.

Liu, H., Jin, L., Koh, S.B.S., Atanasov, I., Schein, S., Wu, L., and Zhou, Z.H. (2010). Atomic Structure of Human Adenovirus by Cryo-EM Reveals Interactions Among Protein Networks. Science *329*, 1038-1043.

Ma, H.-C., and Hearing, P. (2011). Adenovirus Structural Protein IIIa Is Involved in the Serotype Specificity of Viral DNA Packaging. Journal of Virology *85*, 7849-7855.

Mahonen, A.P., Higuchi, M., Tormakangas, K., Miyawaki, K., Pischke, M.S., Sussman, M.R., Helariutta, Y., and Kakimoto, T. (2006). Cytokinins regulate a bidirectional phosphorelay network in Arabidopsis. Current Biology *16*, 1116-1122.

Marsh, M., and Helenius, A. (2006). Virus entry: Open sesame. Cell *124*, 729-740.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proceedings of the National Academy of Sciences of the United States of America *99*, 12246-12251.

Martin, W., Stoebe, B., Goremykin, V., Hapsmann, S., Hasegawa, M., and Kowallik, K.V. (1998). Gene transfer to the nucleus and the evolution of chloroplasts. Nature *393*, 162-165.

Marttila, M., Persson, D., Gustafsson, D., Liszewski, M.K., Atkinson, J.P., Wadell, G., and Arnberg, N. (2005). CD46 is a cellular receptor for all species B adenoviruses except types 3 and 7. Journal of Virology *79*, 14429-14436.

Mascher, T., Helmann, J.D., and Unden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. Microbiology and Molecular Biology Reviews *70*, 910-938.

Meier, O., and Greber, U.F. (2003). Adenovirus endocytosis. Journal of Gene Medicine *5*, 451-462.

Mema, S.C., MacDonald, J., Wyse, J.P.H., Gonder, T., Musto, R., and McIntyre, L. (2010). Public Health adds value to an investigation of epidemic keratoconjunctivitis. Canadian Journal of Ophthalmology-Journal Canadien D Ophtalmologie *45*, 538-539.

Meyer-Ruesenberg, B., Loderstaedt, U., Richard, G., Kaulfers, P.-M., and Gesser, C. (2011). Epidemic Keratoconjunctivitis The Current Situation and Recommendations for Prevention and Treatment. Deutsches Arzteblatt International *108*, 475-480.

Moussatche, P., and Klee, H.J. (2004). Autophosphorylation activity of the Arabidopsis ethylene receptor multigene family. Journal of Biological Chemistry *279*, 48734-48741.

Muller-Dieckmann, H.J., Grantz, A.A., and Kim, S.H. (1999). The structure of the signal receiver domain of the Arabidopsis thaliana ethylene receptor ETR1. Structure with Folding & Design *7*, 1547-1556.

Nemerow, G.R., Pache, L., Reddy, V., and Stewart, P.L. (2009). Insights into adenovirus host cell interactions from structural studies. Virology *384*, 380-388.

Nemerow, G.R., Stewart, P.L., and Reddy, V.S. (2012). Structure of human adenovirus. Current opinion in virology *2*, 115-121.

Nercelles, P.M., Peirano, N.L.N., Herrera, R.O., Rivero, P.B., and Marquez, L.P. (2010). A nosocomial outbreak of epidemic keratoconjunctivitis. Revista Chilena De Infectologia *27*, 534-538.

Neu, U., Bauer, J, and Stehle, T. (2011). Viruses and sialic acids: rules of engagement. Current Opinion in Structural Biology *21*, 610-618.

Neu, U., Stehle, T., and Atwood, W.J. (2009). The Polyomaviridae: Contributions of virus structure to our understanding of virus receptors and infectious entry. Virology *384*, 389-399.

Nilsson, E.C., Storm, R.J., Bauer, J., Johansson, S.M.C., Lookene, A., Angstrom, J., Hedenstrom, M., Eriksson, T.L., Frangsmyr, L., Rinaldi, S*., et al.* (2011). The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. Nature Medicine *17*, 105-109.

Ninfa, A.J. (2010). Use of two-component signal transduction systems in the construction of synthetic genetic networks. Current Opinion in Microbiology *13*, 240-245.

Painter, W., Robertson, A., Trost, L.C., Godkin, S., Lampert, B., and Painter, G. (2012). First Pharmacokinetic and Safety Study in Humans of the Novel Lipid Antiviral Conjugate CMX001, a Broad-Spectrum Oral Drug Active against Double-Stranded DNA Viruses. Antimicrobial Agents and Chemotherapy *56*, 2726-2734.

Paul, R., Jaeger, T., Abel, S., Wiederkehr, I., Folcher, M., Biondi, E.G., Laub, M.T., and Jenal, U. (2008). Allosteric regulation of histidine kinases by their cognate response regulator determines cell fate. Cell *133*, 452-461.

Pekarova, B., Klumpler, T., Triskova, O., Horak, J., Jansen, S., Dopitova, R., Borkovcova, P., Papouskova, V., Nejedla, E., Sklenar, V.*, et al.* (2011). Structure and binding specificity of the receiver domain of sensor histidine kinase CKI1 from Arabidopsis thaliana. Plant Journal *67*, 827-839.

Pelletier, J.S., Stewart, K., Trattler, W., Ritterband, D.C., Braverman, S., Samson, C.M., Liang, B., and Capriotti, J.A. (2009). A combination povidone-iodine 0.4%/dexamethasone 0.1% ophthalmic suspension in the treatment of adenoviral conjunctivitis. Advances in Therapy *26*, 776-783.

Perraud, A.L., Weiss, V., and Gross, R. (1999). Signalling pathways in two-component phosphorelay systems. Trends in Microbiology *7*, 115-120.

Pham, J., Liu, J., Bennett, M.H., Mansfield, J.W., and Desikan, R. (2012). Arabidopsis histidine kinase 5 regulates salt sensitivity and resistance against bacterial and fungal infection. New Phytologist *194*, 168-180.

Pieniazek, N.J., Slemenda, S.B., Pieniazek, D., Velarde, J., and Luftig, R.B. (1990). Human enteric adenovirus type-41 (TAK) contains a 2nd fiber protein gene. Nucleic Acids Research *18*, 1901-1901.

Pischke, M.S., Jones, L.G., Otsuga, D., Fernandez, D.E., Drews, G.N., and Sussman, M.R. (2002). An Arabidopsis histidine kinase is essential for megagametogenesis. Proceedings of the National Academy of Sciences of the United States of America *99*, 15800-15805.

Radtke, K., Dohner, K., and Sodeik, B. (2006). Viral interactions with the cytoskeleton: a hitchhiker's guide to the cell. Cellular Microbiology *8*, 387-400.

Reddy, V.S., Natchiar, S.K., Stewart, P.L., and Nemerow, G.R. (2010). Crystal Structure of Human Adenovirus at 3.5 angstrom Resolution. Science *329*, 1071-1075.

Ricciardi, R.P., Jones, R.L., Cepko, C.L., Sharp, P.A., and Roberts, B.E. (1981). Expression of early adenovirus genes requires a viral encoded acidic polypeptide. Proc Natl Acad Sci U S A *78*, 6121-6125.

Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J.*, et al.* (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature *440*, 940-943.

Rodriguez, F.I., Esch, J.J., Hall, A.E., Binder, B.M., Schaller, G.E., and Bleecker, A.B. (1999). A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. Science *283*, 996-998.

Ruigrok, R.W.H., Barge, A., Albigesrizo, C., and Dayan, S. (1990). Structure of adenovirus fibre: II. Morphology of single fibres. Journal of Molecular Biology *215*, 589-596.

Russell, W.C. (2000). Update on adenovirus and its vectors. Journal of General Virology *81*, 2573-2604.

Russell, W.C. (2009). Adenoviruses: update on structure and function. Journal of General Virology *90*, 1-20.

Safrin, S., Cherrington, J., and Jaffe, H.S. (1997). Clinical uses of cidofovir. Reviews in Medical Virology *7*, 145-156.

Salone, B., Martina, Y., Piersanti, S., Cundari, E., Cherubini, G., Franqueville, L., Failla, C.M., Boulanger, P., and Saggio, I. (2003). Integrin alpha 3 beta 1 is an alternative cellular receptor for adenovirus serotype 5. Journal of Virology *77*, 13448-13454.

San Martin, C. (2012). Latest Insights on Adenovirus Structure and Assembly. Viruses-Basel *4*, 847-877.

Schaller, G.E., Kieber, J.J., and Shiu, S.H. (2008). Two-component signaling elements and histidyl-aspartyl phosphorelays. Arabidopsis Book, 10.1199/tab.0112.

Schaller, G.E., Shiu, S.-H., and Armitage, J.P. (2011). Two-Component Systems and Their Co-Option for Eukaryotic Signal Transduction. Current Biology *21*, R320-R330.

Skerker, J.M., Perchuk, B.S., Siryaporn, A., Lubin, E.A., Ashenberg, O., Goulian, M., and Laub, M.T. (2008). Rewiring the specificity of two-component signal transduction systems. Cell *133*, 1043-1054.

Skevaki, C.L., Galani, I.E., Pararas, M.V., Giannopoulou, K.P., and Tsakris, A. (2011). Treatment of Viral Conjunctivitis with Antiviral Drugs. Drugs *71*, 331-347.

Smith, A.C., Poulin, K.L., and Parks, R.J. (2009). DNA Genome Size Affects the Stability of the Adenovirus Virion. Journal of Virology *83*, 2025-2028.

Spjut, S., Qian, W., Bauer, J., Storm, R., Frangsmyr, L., Stehle, T., Arnberg, N., and Elofsson, M. (2011). A Potent Trivalent Sialic Acid Inhibitor of Adenovirus Type 37 Infection of Human Corneal Cells. Angewandte Chemie-International Edition *50*, 6519-6521.

Strunze, S., Engelke, M.F., Wang, I.H., Puntener, D., Boucke, K., Schleich, S., Way, M., Schoenenberger, P., Burckhardt, C.J., and Greber, U.F. (2011). Kinesin-1-Mediated Capsid Disassembly and Disruption of the Nuclear Pore Complex Promote Virus Infection. Cell Host & Microbe *10*, 210-223.

Symeonidis, N., Jakubowski, A., Pierre-Louis, S., Jaffe, D., Pamer, E., Sepkowitz, K., O'Reilly, R.J., and Papanicolaou, G.A. (2007). Invasive adenoviral infections in T-cell-depleted allogeneic hematopoietic stem cell transplantation: high mortality in the era of cidofovir. Transplant Infectious Disease *9*, 108-113.

Tabbara, K.F., Omar, N., Hammouda, E., Akanuma, M., Ohguchi, T., Ariga, T., Tagawa, Y., Kitaichi, N., Ishida, S., Aoki, K.*, et al.* (2010). Molecular epidemiology of adenoviral keratoconjunctivitis in Saudi Arabia. Molecular Vision *16*, 2132-2136.

Taiz, L., Zeiger, E. (2010). Plant Physiology, Fifth edn (Sinauer Associates Inc.).

Taube, S., Perry, J.W., Yetming, K., Patel, S.P., Auble, H., Shu, L., Nawar, H.F., Lee, C.H., Connell, T.D., Shayman, J.A.*, et al.* (2009). Ganglioside-Linked Terminal Sialic Acid Moieties on Murine Macrophages Function as Attachment Receptors for Murine Noroviruses. Journal of Virology *83*, 4092-4101.

Tollefson, A.E., Ryerse, J.S., Scaria, A., Hermiston, T.W., and Wold, W.S.M. (1996). The E3-11.6-kDa adenovirus death protein (ADP) is required for efficient cell death: Characterization of cells infected with adp mutants. Virology *220*, 152-162.

Tran, L.-S.P., Urao, T., Qin, F., Maruyama, K., Kakimoto, T., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2007). Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America *104*, 20623-20628.

Uchio, E., Fuchigami, A., Kadonosono, K., Hayashi, A., Ishiko, H., Aoki, K., and Ohno, S. (2007). Anti-adenoviral effect of anti-HIV agents in vitro in serotypes inducing keratoconjunctivitis. Graefes Archive for Clinical and Experimental Ophthalmology *245*, 1319-1325.

van Raaij, M.J., Mitraki, A., Lavigne, G., and Cusack, S. (1999). A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature *401*, 935-938.

Varki, A. (2008). Sialic acids in human health and disease. Trends in Molecular Medicine *14*, 351-360.

Viney, K.A., Kehoe, P.J., Doyle, B., Sheppeard, V., Roberts-Witteveen, A.R., Semirli, H., McPhie, K.A., Dwyer, D.E., and McAnulty, J.M. (2008). An outbreak of epidemic keratoconjunctivitis in a regional ophthalmology clinic in New South Wales. Epidemiology and Infection *136*, 1197-1206.

Viswanathan, K., Chandrasekaran, A., Srinivasan, A., Raman, R., Sasisekharan, V., and Sasisekharan, R. (2010). Glycans as receptors for influenza pathogenesis. Glycoconjugate Journal *27*, 561-570.

Voet-van-Vormizeele, J., and Groth, G. (2008). Ethylene controls autophosphorylation of the histidine kinase domain in ethylene receptor ETR1. Molecular plant *1*, 380-387.

Walters, R.W., Freimuth, P., Moninger, T.O., Ganske, I., Zabner, J., and Welsh, M.J. (2002). Adenovirus fiber disrupts CAR-mediated intercellular adhesion allowing virus escape. Cell *110*, 789-799.

Wang, H., Li, Z.-Y., Liu, Y., Persson, J., Beyer, I., Moeller, T., Koyuncu, D., Drescher, M.R., Strauss, R., Zhang, X.-B.*, et al.* (2011). Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. Nature Medicine *17*, 96-104.

Wickham, T.J., Mathias, P., Cheresh, D.A., and Nemerow, G.R. (1993). Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. Cell *73*, 309-319.

Wiethoff, C.M., Wodrich, H., Gerace, L., and Nemerow, G.R. (2005). Adenovirus protein VI mediates membrane disruption following capsid disassembly. Journal of Virology *79*, 1992-2000.

Wohlbach, D.J., Quirino, B.F., and Sussman, M.R. (2008). Analysis of the Arabidopsis histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation. Plant Cell *20*, 1101-1117.

Wuichet, K., Cantwell, B.J., and Zhulin, I.B. (2010). Evolution and phyletic distribution of two-component signal transduction systems. Current Opinion in Microbiology *13*, 219-225.

Wulfetange, K., Lomin, S.N., Romanov, G.A., Stolz, A., Heyl, A., and Schmuelling, T. (2011). The Cytokinin Receptors of Arabidopsis Are Located Mainly to the Endoplasmic Reticulum. Plant Physiology *156*, 1808-1818.

Yoo, S.-D., Cho, Y.-H., Tena, G., Xiong, Y., and Sheen, J. (2008). Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signalling. Nature *451*, 789-795.

Zhou, X.H., Robinson, C.M., Rajaiya, J., Dehghan, S., Seto, D., Jones, M.S., Dyer, D.W., and Chodosh, J. (2012). Analysis of Human Adenovirus Type 19 Associated with Epidemic Keratoconjunctivitis and its Reclassification as Adenovirus Type 64. Invest Ophthalmol Vis Sci *53*, 2804-2811.

8 Appendix

8.1 Acknowledgments

8.2 Publications

Nilsson, E.C., Storm, R.J., **Bauer, J.**, Johansson, S.M.C., Lookene, A., Angstrom, J., Hedenstrom, M., Eriksson, T.L., Frangsmyr, L., Rinaldi, S*., et al.* **(2011)**. The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. Nature Medicine *17*, 105-109.

Spjut, S., Qian, W., **Bauer, J.**, Storm, R., Frangsmyr, L., Stehle, T., Arnberg, N., and Elofsson, M. **(2011)**. A Potent Trivalent Sialic Acid Inhibitor of Adenovirus Type 37 Infection of Human Corneal Cells. Angewandte Chemie International Edition *50*, 6519-6521.

Neu, U., **Bauer, J.**, and Stehle, T. **(2011)**. Viruses and sialic acids: rules of engagement. Current Opinion in Structural Biology *21*, 610-618.

Bauer, J., Reiss, K., Veerabagu, M., Heunemann, M., Harter, K., and Stehle, T. **(2012)**. Structure-function analysis of Arabidopsis thaliana histidine kinase AHK5 bound to its cognate phosphotransfer protein AHP1. Mol Plant, doi: 10.1093/mp/sss126.

8.3 Contributions

8.3.1 Contributions to Nilsson et al.:

I carried out expression, purification and crystallization of the Ad37 fiber knob protein. Following crystallization, I soaked native crystals with the GD1a hexasaccharide, followed by data collection, data processing, structure determination, and structural refinement. For the manuscript, I wrote the description of the structural analysis of the Ad37-GD1a crystal structure and the respective part of the "Methods"-section. In addition, I prepared figures 3A-D, supplemental figures 4F-H, and supplemental table 5.

8.3.2 Contributions to Spjut et al.:

I carried out expression and purification of the Ad37 fiber knob protein. Following purification, I co-crystallized the Ad37 fiber knob together with conjugate ME0322, followed by data collection, data processing, structure determination, and structural refinement. For the manuscript, I wrote the description of the structural analysis of the Ad37-ME0322 crystal structure, and I wrote the paragraph "Protein production and structure determination" in the "Methods"-section. In addition, I prepared figure 2, supplemental figure S1, and supplemental table S1.

8.3.3 Contributions to Neu et al.:

For this review, I prepared Figure 4 and wrote the paragraph "Achieving specificity through multivalent binding".

8.3.4 Contributions to Bauer et al.:

Klaus Harter provided the cDNA of AHK5 and AHP1-6. I conducted construct design and cloning of $AHK5_{RD}$ and AHP1, AHP3, AHP4, AHP5, and AHP6 into the respective expression plasmids. I carried out expression and purification of $AHK5_{RD}$ and AHP1-6, and performed X-ray crystallographic studies. Kerstin Reiss and I conducted SPR experiments. Kerstin Reiss analyzed the data, wrote the respective section in the manuscript and in the "Methods"-section. Moreover, Kerstin Reiss prepared figure 3A-F and table 2. Manikandan Veerabagu, Michael Heunemann and Klaus Harter designed and carried out BiFC experiments, and analyzed the resulting data. In addition, they contributed the BiFC analysis part in the manuscript and in the "Methods"-section, and prepared figure 3G, H. I prepared all other

figures and tables, and I wrote the bulk of the manuscript, together with Thilo Stehle and Klaus Harter.

nature
medicine

The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis

Emma C Nilsson[1], Rickard J Storm[1], Johannes Bauer[2], Susanne M C Johansson[1], Aivar Lookene[3], Jonas Ångström[4], Mattias Hedenström[5], Therese L Eriksson[1], Lars Frängsmyr[1], Simon Rinaldi[6], Hugh J Willison[6], Fatima Pedrosa Domellöf[7,8], Thilo Stehle[2,9] & Niklas Arnberg[1,10]

Adenovirus type 37 (Ad37) is a leading cause of epidemic keratoconjunctivitis (EKC)[1,2], a severe and highly contagious ocular disease. Whereas most other adenoviruses infect cells by engaging CD46 or the coxsackie and adenovirus receptor (CAR), Ad37 binds previously unknown sialic acid–containing cell surface molecules[3,4]. By glycan array screening, we show here that the receptor-recognizing knob domain of the Ad37 fiber protein specifically binds a branched hexasaccharide that is present in the GD1a ganglioside and that features two terminal sialic acids. Soluble GD1a glycan and GD1a-binding antibodies efficiently prevented Ad37 virions from binding and infecting corneal cells. Unexpectedly, the receptor is constituted by one or more glycoproteins containing the GD1a glycan motif rather than the ganglioside itself, as shown by binding, infection and flow cytometry experiments. Molecular modeling, nuclear magnetic resonance and X-ray crystallography reveal that the two terminal sialic acids dock into two of three previously established sialic acid–binding sites in the trimeric Ad37 knob. Surface plasmon resonance analysis shows that the knob–GD1a glycan interaction has high affinity. Our findings therefore form a basis for the design and development of sialic acid–containing antiviral drugs for topical treatment of EKC.

EKC is caused mainly by three species D adenoviruses (Ad8, Ad19 and Ad37) and is recognized as a severe ocular disease for which no antiviral drugs are available[2,5,6]. Common symptoms of EKC are pain, edema, lacrimation and decreased vision that can last for months or years. It has been suggested that two receptors, CD46 and α2,3-linked sialic acid, are used by EKC-causing adenoviruses (**Supplementary Fig. 1** and refs. 3,4,7–9), but only sialic acid has been confirmed to serve as a functional cellular receptor[10–12]. Other (non–EKC-causing) species D adenoviruses also bind cell surface

sialic acid, but it is unclear whether this leads to productive infection[7,13]. To identify the exact structure of the sialic acid–containing glycans used by EKC-causing adenoviruses, we analyzed binding of Ad37 knob proteins in a glycan array. The knob bound a number of sialylated, soluble plasma proteins (peak nos. 1–6) that we did not consider further, sulfated glycans (nos. 26–40 and 42–47), and, notably, glycan no. 212, which corresponds to the disialylated glycan found in ganglioside GD1a (**Fig. 1** and **Supplementary Table 1**). Previous studies of Ad37 showed that sulfated glycans serve only as nonfunctional receptors[7], and, in agreement, we found that soluble, sulfated disaccharides did not prevent Ad37 virion binding to cells (**Supplementary Table 2**). To test whether the GD1a glycan could serve as a functional Ad37 receptor, we preincubated virions with soluble GD1a glycan and found that this efficiently inhibited binding and infection of human corneal epithelial (HCE) cells compared to control glycans (**Fig. 2a,b**). Moreover, preincubation of cells with a monoclonal antibody (clone EM9) recognizing GD1a efficiently inhibited binding and infection compared to control antibodies (**Fig. 2c,d**; for specificity of EM9 see **Supplementary Fig. 2a–c**). Finally, binding of the Ad37 knob to HCE cells was outcompeted by soluble GD1a glycan, but not by control glycans (**Fig. 2e**) at the concentrations used, and immunohistochemistry experiments confirmed that both the Ad37 knob and the EM9 antibody bound the surface of all layers of primary HCE cells *in situ* (**Fig. 2f–i**). Taken together, these results demonstrate that Ad37 virions bind corneal cell surface molecules containing the GD1a glycan motif.

To determine whether Ad37 binds the GD1a ganglioside itself, we first screened Ad37 knob binding to a combinatorial glycolipid array[14] containing GD1a and several other gangliosides, and found no binding (data not shown). Additionally, no binding of soluble Ad37 knob to immobilized GD1a or other gangliosides could be detected by ELISA (data not shown). Next, we analyzed Ad37 virion binding to sialic acid–deficient Lec2 cells to which GD1a gangliosides had either been

[1]Department of Clinical Microbiology, Division of Virology, Umeå University, Umeå, Sweden. [2]Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. [3]Department of Chemistry, Tallinn University of Technology, Tallinn, Estonia. [4]Institute of Biomedicine, Department of Medical Biochemistry and Cell Biology, University of Göteborg, Göteborg, Sweden. [5]Department of Chemistry, Computational Life Science Cluster (CLiC), Umeå University, Umeå, Sweden. [6]Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, Glasgow Biomedical Research Centre, University of Glasgow, Glasgow, UK. [7]Department of Clinical Sciences, Ophthalmology, Umeå University, Umeå, Sweden. [8]Department of Integrative Medical Biology, Anatomy, Umeå University, Umeå, Sweden. [9]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [10]Laboratory for Molecular Infection Medicine in Sweden (MIMS), Umeå University, Umeå, Sweden. Correspondence should be addressed to N.A. (niklas.arnberg@climi.umu.se).

Figure 1 Glycan array of Ad37 knob interactions with 260 different glycans. Elevated peaks indicate that the Ad37 knob bound with high efficiency to immobilized plasma glycoproteins (glycan nos. 1–6), sulfated glycans (nos. 26–40 and 42–47), monosialic acid–containing glycans (nos. 14–16), and one branched, disialic acid–containing glycan (no. 212) corresponding to the hexasaccharide that is found in ganglioside GD1a (see **Supplementary Table 1** for details). RFU, relative fluorescent units.
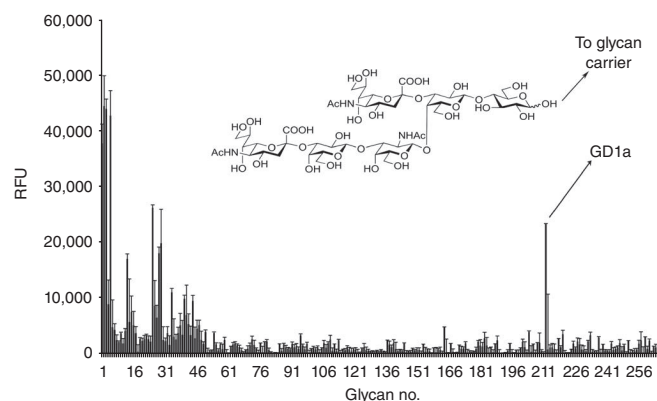
added or not (control), but found no substantially increased binding compared to control cells (**Supplementary Fig. 3a**), despite efficient incorporation of GD1a gangliosides into Lec2 cell membranes, as determined by flow cytometry (**Supplementary Fig. 3b**). Furthermore, metabolic reduction of *de novo* ganglioside biosynthesis did not affect Ad37 binding (**Supplementary Fig. 3c,d**), and GD1a ganglioside–containing liposomes did not compete with binding of Ad37 virions to HCE cells (**Supplementary Fig. 3e**). We next tested whether the Ad37 receptor instead contains a protein component. Accordingly, binding of HCE cells by Ad37 virions was efficiently reduced by enzymatic removal of cell surface proteins (**Supplementary Fig. 3f**), and inhibition of *de novo* synthesis of O-linked (via Ser or Thr) glycans by pretreating HCE cells with benzyl-α-*N*-acetylgalactosamine (benzyl-α-GalNAc) efficiently prevented Ad37 binding and infection of HCE cells (**Supplementary Fig. 3g,h**). However, removal of cell surface N-glycans by treatment with N-glycosidase F (PNGase F) did not inhibit Ad37 binding (data not shown). Moreover, Ad37 knob overlay experiments showed that the knobs bound several HCE cell membrane proteins (data not shown), although these studies are inconclusive in that one cannot exclude the possibility of false positives (i.e., binding to intracellular proteins). Nevertheless, our results suggest that Ad37 binds one or more cell surface glycoproteins rather than to the GD1a ganglioside itself.

To understand the determinants of the specific Ad37 knob–GD1a interaction revealed by the glycan array screen, we first modeled the glycan part of GD1a onto the Ad37 knob. By systematic variation of the glycosidic torsion angles representing known energy minima for

the GD1a glycan and constituent disaccharide units, the two sialic acids of GD1a can be docked into two of the three previously established sialic acid binding sites in the trimeric Ad37 knob (**Supplementary Fig. 4a–c**). To further investigate the binding epitope of the GD1a glycan, we carried out saturation transfer difference (STD) NMR experiments[15]. The largest effect is seen in the methyl peaks from the acetamide groups from the two sialic acids in the GD1a glycan, *N*-acetylneuraminic acids A and B (Neu5Ac^A and Neu5Ac^B) (**Supplementary Fig. 4d**), indicating that these form contacts with Ad37. Although these peaks overlap for Neu5Ac^A and Neu5Ac^B, it is clear that both sialic acids participate in binding, as is seen on the H3eq and H3ax protons that are resolved for Neu5Ac^A and Neu5Ac^B (**Supplementary Fig. 4e**). Although not all peaks in the STD spectrum could be unambiguously assigned, the most prominent peaks all belonged to Neu5Ac^A and Neu5Ac^B, and we detected little, if any, effect from remaining GD1a sugar residues, as exemplified by Gal^{II}H2 and Glc^{I}H2 (**Supplementary Fig. 4e**). A control STD experiment (**Supplementary Fig. 4d**, bottom spectrum) in the absence of protein showed that a negligible amount of direct saturation of the GD1a glycan occurs with the experimental parameters used here. No peaks were detected from galactose that was added as
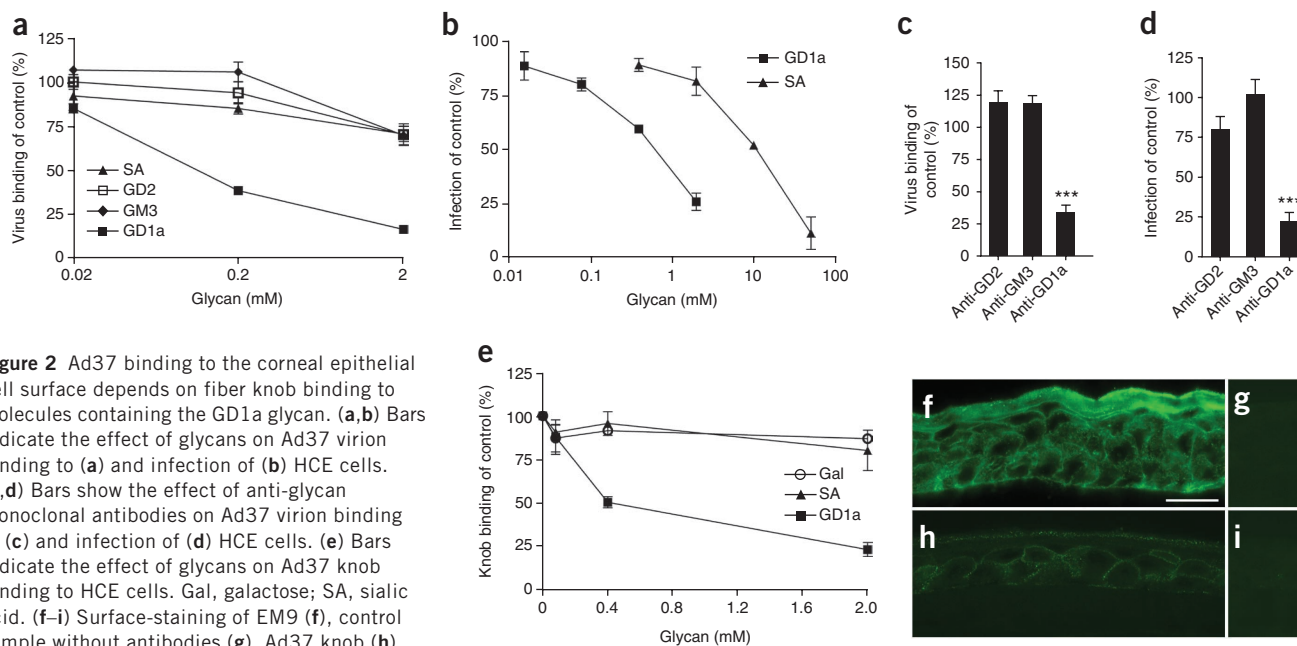

Figure 2 Ad37 binding to the corneal epithelial cell surface depends on fiber knob binding to molecules containing the GD1a glycan. (**a,b**) Bars indicate the effect of glycans on Ad37 virion binding to (**a**) and infection of (**b**) HCE cells. (**c,d**) Bars show the effect of anti-glycan monoclonal antibodies on Ad37 virion binding to (**c**) and infection of (**d**) HCE cells. (**e**) Bars indicate the effect of glycans on Ad37 knob binding to HCE cells. Gal, galactose; SA, sialic acid. (**f–i**) Surface-staining of EM9 (**f**), control sample without antibodies (**g**), Ad37 knob (**h**), and control without Ad37 (**i**) for all layers of human corneal epithelium. Scale bar, 20 μm. Controls were performed without glycans (**a**, **b**, **e**); without antibodies (**c**, **d**, **g**) and without Ad37 knob protein (**e**, **i**). The control values were set to 100% on the *y* axes in **a**–**e**. Error bars in **a**–**e** show ± s.e.m.
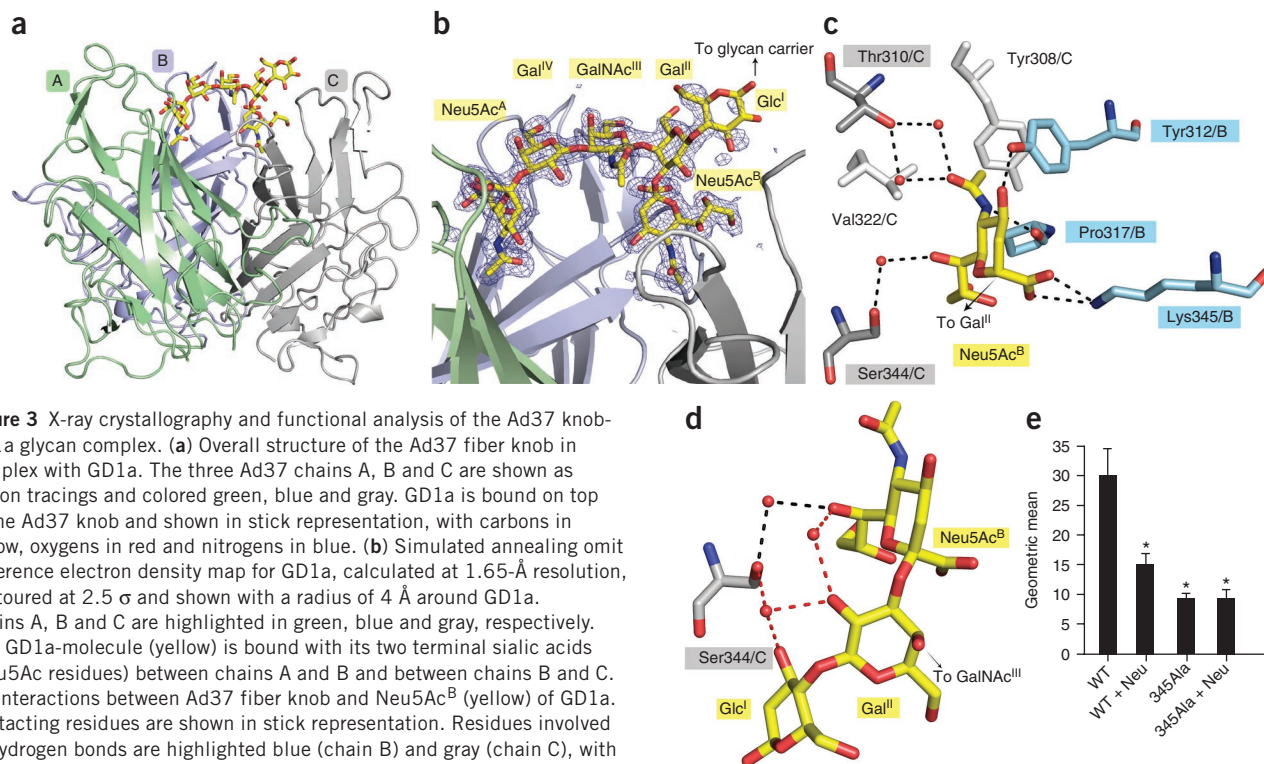
Figure 3 X-ray crystallography and functional analysis of the Ad37 knob-GD1a glycan complex. (**a**) Overall structure of the Ad37 fiber knob in complex with GD1a. The three Ad37 chains A, B and C are shown as ribbon tracings and colored green, blue and gray. GD1a is bound on top of the Ad37 knob and shown in stick representation, with carbons in yellow, oxygens in red and nitrogens in blue. (**b**) Simulated annealing omit difference electron density map for GD1a, calculated at 1.65-Å resolution, contoured at 2.5 σ and shown with a radius of 4 Å around GD1a. Chains A, B and C are highlighted in green, blue and gray, respectively. One GD1a-molecule (yellow) is bound with its two terminal sialic acids (Neu5Ac residues) between chains A and B and between chains B and C. (**c**) Interactions between Ad37 fiber knob and Neu5Ac^B (yellow) of GD1a. Contacting residues are shown in stick representation. Residues involved in hydrogen bonds are highlighted blue (chain B) and gray (chain C), with oxygens in red and nitrogens in blue. Also shown: hydrogen bonds (broken lines), water molecules (red spheres) and residues only involved in van der Waals contacts (light gray). (**d**) Hydrogen bond network between Ad37 fiber knob and GD1a. Hydrogen bonds to the bridging sugar residues Gal^{II} and Glc^I are shown as red broken lines. Previously introduced hydrogen bonds (shown in **c**) between Ad37 knob and Neu5Ac^B are indicated by black broken lines; water molecules are shown as red spheres. (**e**) Bars indicate flow cytometry analysis of knob binding to HCE cells treated with or without neuraminidase. WT, wild-type. Error bars show ± s.e.m.

a negative control (**Supplementary Fig. 4e**, top spectrum). All peaks in the STD spectrum are therefore directly related to binding of the GD1a glycan to Ad37.

To verify the structural features of the Ad37 knob–GD1a glycan interaction, we solved the crystal structure of the Ad37 knob complexed with GD1a oligosaccharide at 1.65-Å resolution. The knob engages a single GD1a glycan. The entire GD1a oligosaccharide is clearly visible in the electron density map, allowing its unambiguous placement. The two terminal sialic acids make direct interactions with the Ad37 knob (**Fig. 3a,b**) and bind very similarly to two different protomers in the Ad37 knob (**Fig. 3c** and **Supplementary Fig. 4b,c**). In each case, direct interactions involve Tyr312, Pro317 and Lys345 of chains A and B, respectively (**Fig. 3c**). Water-mediated hydrogen bonds to Thr310 and Ser344 provide additional contacts. Both galactose residues interact through one water molecule with the Ad37 knob (**Fig. 3d**). For the most distal galactose (Gal^{IV}), the C6 hydroxyl group binds Ser344 in chain B (Ser344/B; data not shown), whereas the other galactose (Gal^{II}) forms a hydrogen bond with its C2 hydroxyl group to Ser344/C. The glucose forms a water-mediated hydrogen bond with its C3 hydroxyl group to Ser344/C, and GalNAc^{III} forms one water-mediated hydrogen bond with its C3 hydroxyl group to the amino group of Lys345/B (not shown). A crucial role of Lys345 was suggested by the *in silico* substitution Lys345Ala (**Supplementary Table 3**). Flow cytometry experiments with wild-type Ad37 knob and the Lys345Ala mutant confirmed this by showing that the mutant knob had lost most of the sialic acid–dependent cell-binding capacity (**Fig. 3e**). Only two sialic acid binding sites are accessible for the GD1a glycan in the crystal (**Supplementary Fig. 4f–h**). The third binding site is sterically blocked by a symmetry-related molecule, thus preventing

the binding of either sialylated arm of the GD1a glycan to this pocket. In solution, however, all three binding sites are accessible, and there is no preference for any of the three binding sites.

We further analyzed the affinity and kinetics of GD1a glycan binding to the Ad37 knob with surface plasmon resonance (SPR) (**Supplementary Fig. 5a**). The association rate constant determined at low GD1a glycan concentrations was more than ten times higher than the constant obtained at high concentrations (**Supplementary Table 4**), whereas the dissociation rate constant was not greatly dependent on the concentrations of the glycan. Our data therefore suggest that the GD1a glycan binds at least two different regions of the knob protein:
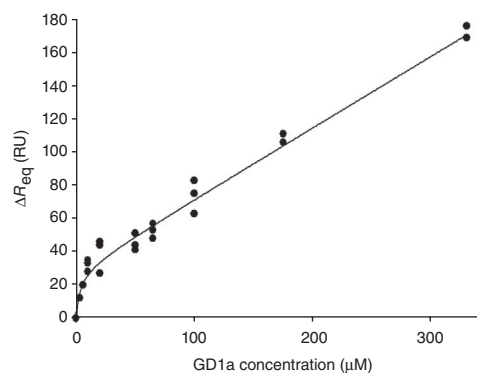

Figure 4 Surface plasmon resonance analysis of GD1a glycan binding to immobilized Ad37 knob protein. Response values at equilibrium, ΔR_{eq}, are shown for each GD1a concentration. The curve was obtained by fitting the data to a two–binding site model.

a high-affinity (K_d = 19 μM) binding site detected at low glycan concentrations (3–10 μM) and a lower-affinity (K_d = 265 μM) binding site observed only at higher glycan concentrations (100–300 μM). This binding mode is supported by equilibrium analysis of the data: when response values at equilibrium (ΔR_{eq}) were plotted against the injected glycan concentrations, the result was not a typical saturable hyperbola (**Fig. 4**). The ΔR_{eq} values increased more steeply at low concentrations of the glycan, but they increased less steeply and almost linearly at higher concentrations. The more than 250-fold-higher affinity (19 μM) described here for the GD1a glycan–Ad37 knob interaction compared with the sialyllactose–Ad37 knob interaction (K_d = 5 mM)[16] is probably due to a single GD1a glycan interacting directly with the knob through both sialic acids. The ΔR_{eq} value increased more at high glycan concentrations (the linear part of the curve) than at low concentrations. This indicates that there are more low-affinity binding sites than high-affinity binding sites for the glycan at the surface of the knob, suggesting that the low-affinity interaction is not mediated by one sialic acid of the GD1a glycan binding to the remaining free sialic acid binding site (at high concentrations of the GD1a glycan, the other two sialic acid binding sites are occupied by the two sialic acids in one GD1a glycan). In support of this, the low-affinity interaction was not dissociated in the presence of 1M NaCl (**Supplementary Fig. 5b**), in contrast to the high-affinity interaction (**Supplementary Fig. 5c**). Thus, the low-affinity interaction needs further evaluation.

Other adenoviruses are capable of engaging three separate receptor molecules (CAR or CD46) per fiber knob with high affinity[17–19], leading to subsequent internalization. As the Ad37 knob only engages one GD1a receptor molecule, receptor clustering around one single fiber knob does not appear necessary for subsequent internalization of Ad37. The affinity of the Ad37 knob–GD1a glycan interaction (K_d = 19 μM) is lower than those of the Ad-CAR or Ad-CD46 interactions[18,19] but much higher than the interaction of the influenza A virus hemagglutinin with sialic acid (K_d in mM range)[20]. Combined with the efficient inhibition of Ad37 infection by GD1a glycan–binding antibodies, this suggests that Ad37 engagement of GD1a-containing cell surface glycans is specific and leads to subsequent internalization.

The GD1a glycan–contacting residues are conserved in many species D adenoviruses, including the three that cause EKC (**Supplementary Fig. 6a**), suggesting that many adenovirus types could interact with the GD1a glycan. In support of this suggestion, flow cytometry analyses show that knobs from both EKC-causing types (Ad8, Ad19a and Ad37) and non–EKC-causing types (Ad9 and Ad19p) bind less efficiently to HCE cells in the presence of soluble GD1a glycan (**Supplementary Fig. 6b**). Ad5 (species C) binding was not affected by the GD1a glycan. Finally, *in silico* substitution of the residues distinguishing the Ad37 knob from those of Ad19p and Ad8 also supports the notion that multiple species D adenoviruses are capable of interacting with the GD1a glycan (**Supplementary Table 3**). The knobs of Ad19a and Ad37 are identical and differ from Ad19p at two residues only[21]. The inability of Ad19p and other species D adenoviruses (except Ad8, Ad19a and Ad37) to cause EKC is therefore likely to depend on factors other than their GD1a glycan–binding capacity. Whereas this study shows that the EKC-causing adenoviruses can use the GD1a glycan as a cellular receptor, further studies are required to elucidate the role of this molecule as a receptor for the large number of other, non-EKC-causing species D adenoviruses (>30 different types).

Currently there are no antiviral drugs available to treat EKC, but such drugs are highly desirable both for health and socioeconomic reasons[1,6,22]. The discovery of the GD1a glycan as a cellular receptor for Ad37 and the detailed structural and functional analysis of the interaction provide a useful platform for design and development of small molecule entry inhibitors. Compounds modeled on sialic acid have proven to be highly useful for antiviral therapy in the case of influenza[23,24]. The main difference between Ad37 knob–GD1a glycan interaction and influenza hemagglutinin interaction with sialoglycans is that the sialic acid binding sites lie at the tip of the hemagglutinin protein, whereas on the knob they lie around the central cavity. For treatment of EKC, sialic acid–based drugs would preferably be given topically and would function by preventing newly formed virions from binding to noninfected cells; this would thereby limit symptoms, shorten duration of illness and prevent transmission to noninfected individuals. Such a drug could even form a template for the development of molecules that target other sialic acid–binding viruses, such as influenza A viruses.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemedicine/.

Note: Supplementary information is available on the Nature Medicine website.

AUTHOR CONTRIBUTIONS
E.C.N. and R.J.S. contributed equally to design and conduction of binding, infection and flow cytometry experiments; S.M.C.J. produced GD1a glycan and performed NMR studies with M.H.; J.B. and T.S. carried out X-ray crystallography studies; S.R. and H.J.W. performed combinatorial glycolipid glycoarray; J.Å. conducted molecular modeling; F.P.D. carried out immunohistochemistry analysis; and A.L. performed SPR experiments. L.F. and T.L.E. conducted two-dimensional gel electrophoresis and blotting experiments, and L.F. did statistical calculations. T.S., M.H., F.P.D., A.L., J.Å., H.J.W., L.F., J.B. and N.A. discussed and wrote the manuscript, and T.S. and N.A. supervised the project.

1. Wold, W.S.M. & Horwitz, M.S. Adenoviruses. in *Fields Virology*, Vol. 2 (eds. Knipe, D.M. & Howley, P.M.) 2395–2436. (Lippincott Williams & Wilkins, Philadelphia, 2007).
2. Ford, E., Nelson, K.E. & Warren, D. Epidemiology of epidemic keratoconjunctivitis. *Epidemiol. Rev.* **9**, 244–261 (1987).
3. Arnberg, N. Adenovirus receptors, implications for tropism, treatment and targeting. *Rev. Med. Virol.* **19**, 165–178 (2009).

4. Arnberg, N., Edlund, K., Kidd, A.H. & Wadell, G. Adenovirus type 37 uses sialic acid as a cellular receptor. *J. Virol.* **74**, 42–48 (2000).
5. Gordon, Y.J., Aoki, K. & Kinchington, P.R. Adenovirus keratoconjunctivitis. in *Ocular Infection and Immunity* (eds. Pepose, J.S., Holland, G.N. & Wilhelmus, K.R.) 877–894. (Mosby, St. Louis, 1996).
6. Kinchington, P.R., Romanowski, E.G. & Gordon, Y.J. Prospects for adenovirus antivirals. *J. Antimicrob. Chemother.* **55**, 424–429 (2005).
7. Arnberg, N. *et al.* Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction. *Virology* **302**, 33–43 (2002).
8. Arnberg, N., Pring-Akerblom, P. & Wadell, G. Adenovirus type 37 uses sialic acid as a cellular receptor on Chang C cells. *J. Virol.* **76**, 8834–8841 (2002).
9. Wu, E. *et al.* Membrane cofactor protein is a receptor for adenoviruses associated with epidemic keratoconjunctivitis. *J. Virol.* **78**, 3897–3905 (2004).
10. Cashman, S.M., Morris, D.J. & Kumar-Singh, R. Adenovirus type 5 pseudotyped with adenovirus type 37 fiber uses sialic acid as a cellular receptor. *Virology* **324**, 129–139 (2004).
11. Lecollinet, S. *et al.* Improved gene delivery to intestinal mucosa by adenoviral vectors bearing subgroup B and D fibers. *J. Virol.* **80**, 2747–2759 (2006).
12. Thirion, C. *et al.* Adenovirus vectors based on human adenovirus type 19a have high potential for human muscle-directed gene therapy. *Hum. Gene Ther.* **17**, 193–205 (2006).
13. Arnberg, N., Kidd, A.H., Edlund, K., Olfat, F. & Wadell, G. Initial interactions of subgenus D adenoviruses with A549 cellular receptors: sialic acid versus α_v integrins. *J. Virol.* **74**, 7691–7693 (2000).
14. Rinaldi, S. *et al.* Analysis of lectin binding to glycolipid complexes using combinatorial glycoarrays. *Glycobiology* **19**, 789–796 (2009).
15. Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* **38**, 1784–1788 (1999).
16. Burmeister, W.P., Guilligay, D., Cusack, S., Wadell, G. & Arnberg, N. Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J. Virol.* **78**, 7727–7736 (2004).
17. Bewley, M.C., Springer, K., Zhang, Y.B., Freimuth, P. & Flanagan, J.M. Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. *Science* **286**, 1579–1583 (1999).
18. Persson, B.D. *et al.* Adenovirus type 11 binding alters the conformation of its receptor CD46. *Nat. Struct. Mol. Biol.* **14**, 164–166 (2007).
19. Kirby, I. *et al.* Identification of contact residues and definition of the CAR-binding site of adenovirus type 5 fiber protein. *J. Virol.* **74**, 2804–2813 (2000).
20. Sauter, N.K. *et al.* Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. *Biochemistry* **28**, 8388–8396 (1989).
21. Arnberg, N., Mei, Y. & Wadell, G. Fiber genes of adenoviruses with tropism for the eye and the genital tract. *Virology* **227**, 239–244 (1997).
22. Lenaerts, L., De Clercq, E. & Naesens, L. Clinical features and treatment of adenovirus infections. *Rev. Med. Virol.* **18**, 357–374 (2008).
23. von Itzstein, M. *et al.* Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* **363**, 418–423 (1993).
24. Kim, C.U. *et al.* Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **119**, 681–690 (1997).

ONLINE METHODS

Glycan array. Glycan array was performed as described here: (http://www.functionalglycomics.org/glycomics/HServlet?operation=view&sideMenu=no&psId=primscreen_PA_v2_409_07312006).

Biological competition experiments. Binding, infection and flow cytometry experiments were performed essentially as described previously[7,25], but they were different in that here we used HCE cells[26] (provided by K. Araki-Sasaki) or sialic acid–deficient Lec2 (ref. 4) cells (ATCC), and modified experiments as described below.

For binding experiments, we used 1×10^5 cells and 5,000 virions per cell. HCE cells were treated with or without 25 mU α2,3-specific neuraminidase (Sigma-Aldrich) before incubation with CsCl-purified, ^{35}S-labeled virions in suspension (**Supplementary Fig. 1a**), as previously described[7]. Lec2 cells were preincorporated with (or without) 50 µM GD1a gangliosides (provided by F. Lindh) before incubation with ^{35}S-labeled virions in suspension (**Supplementary Fig. 3a**) as previously described[27]. Virions were preincubated with GD1a glycans, sialic acid (N-acetylneuraminic acid; Sigma-Aldrich), or GM3 or GD2 glycans (both provided by the Consortium for Functional Glycomics) as indicated (**Fig. 2a**), or with 5.7 mM GD1a ganglioside–containing liposomes (provided by A. Carlsson) before incubation with cells in suspension (**Supplementary Fig. 3e**). Liposomes were made in PBS and contained 10 mol% GD1a ganglioside and 90 mol% dioleyl phosphatidyl choline. Liposomes were extruded 11 times through a 100 nM membrane and diluted 1:1 in binding buffer[7,25]. HCE cells (in suspension) were preincubated with 20 µg ml[−1] of antibodies to GD1a (clone EM9)[28], GM3 (clone M2590, Cosmobio) or GD2 (clone 4G12, Abcam) and then incubated with virions (**Fig. 2c**). HCE cells (in suspension) were treated with increasing concentrations of V8 or ficin proteases (Sigma-Aldrich)[4] (**Supplementary Fig. 3f**), or, adherent HCE cells were first incubated with 2.5 µM P4 compounds (provided by R. Schnaar) at 37 °C for 3 d (**Supplementary Fig. 3c**) or with 3 mM benzyl-α-GalNAc (Sigma-Aldrich) at 37 °C for 48 h (**Supplementary Fig. 3g**) before incubation with virions (in suspension). Bound virions were quantified with a scintillation counter.

In the infection experiments, Ad37 virions were preincubated with glycans (**Fig. 2b**) or HCE cells were preincubated with antibodies (20 µg ml[−1]; **Fig. 2d**), with benzyl-α-GalNAc (3 mM; **Supplementary Fig. 3h**) or with 25 mU α2,3-specific neuraminidase (**Supplementary Fig. 1b**) as previously described[7], before infection, and then analyzed by fluorescence microscopy after 44 h.

If not otherwise stated in figure legends, the flow cytometry experiments (**Figs. 2e** and **3e** and **Supplementary Fig. 6b**) were conducted as previously described[7], by preincubating RGS(6)His-tagged knobs with glycans and then with 10[5] HCE cells per sample. Ad37 knob with the Lys345Ala mutation was produced with a QuikChange mutagenesis kit (Agilent). Detection of GD1a ganglioside incorporation in Lec2 cells (**Supplementary Fig. 3b**) was performed by incubating Lec2 cells in 100 µl PBS + 2% (vol/vol) FCS supplemented with 0.5 µg EM9 antibody and FITC-labeled goat antibody to mouse IgM (Dako) diluted 1 in 100. Detection of GM1 gangliosides on HCE cells (**Supplementary Fig. 3d**) after treatment with or without 2.5 µM P4 compounds at 37 °C for 3 d was performed by incubating HCE cells with 1 µg FITC-labeled cholera toxin B (Invitrogen) in 100 µl PBS.

X-ray crystallography: expression and purification. Ad37 knobs were expressed and purified as described previously[16] with the following modifications. We used BL21(DE3)pLysS cells (Novagen) for expression, and after cell lysis and centrifugation the supernatant was loaded onto a 1-ml HisTrap HP column (GE Life Sciences). Washing steps with 100 mM and 200 mM imidazole in the loading buffer were followed by protein elution with 300 mM imidazole. Polyhistidine-tag cleavage was followed by gel filtration (Superdex 200; GE Life Sciences) with 30 mM Tris (pH 7.5) and 150 mM NaCl. Ad37 knob fractions were pooled and applied onto a HisTrap HP column for repurification. Cleaved Ad37 knob was concentrated to 15.0 mg ml[−1].

X-ray crystallography: crystallization and structure determination. Crystals of Ad37 knobs were grown at 20 °C by the hanging drop method using a reservoir solution of 26% (wt/vol) polyethylene glycol 8000, 50 mM zinc acetate, and 100 mM HEPES (pH 7.1). For complex production, crystals were soaked for 2h in reservoir solution containing 10 mM GD1a glycan. Crystals were then frozen in liquid nitrogen and used to collect data at beamline MX 14-2 (wavelength λ = 0.91841 Å) at BESSY II (Berlin, Germany). Diffraction data were recorded with a MX-225 charge-coupled device detector and processed with XDS software[29]. The structure was solved by molecular replacement with Phaser[30] in CCP4[31] and the native Ad37 knob trimer (Protein Data Bank code: 1uxe[16]) as the search model. GD1a was unambiguously located in $F_{soaked} - F_{native}$ difference Fourier maps, incorporated into the model, and refined with restraints from the Refmac5 (ref. 32) monomer library. Refinement was carried out by alternating rounds of model building in Coot[33] and restrained refinement including the translation-libration-screw method (TLS) with Refmac5. For TLS refinement, each protomer in the asymmetric unit (and the water chain) was attributed to one TLS group. Simulated annealing was carried out with PHENIX[34]. The final model had excellent geometry. The asymmetric unit contained one Ad37 fiber knob trimer and one GD1a glycan ligand. All figures were prepared with PyMOL (http://www.pymol.org). Statistics for data collection and refinement are given in **Supplementary Table 5**.

Additional methods. Detailed methodology is described in the **Supplementary Methods**.

25. Johansson, S.M. et al. Multivalent sialic acid conjugates inhibit adenovirus type 37 from binding to and infecting human corneal epithelial cells. *Antiviral Res.* **73**, 92–100 (2007).
26. Araki-Sasaki, K. et al. An SV40-immortalized human corneal epithelial cell line and its characterization. *Invest. Ophthalmol. Vis. Sci.* **36**, 614–621 (1995).
27. Tsai, B. et al. Gangliosides are receptors for murine polyoma virus and SV40. *EMBO J.* **22**, 4346–4355 (2003).
28. Boffey, J. et al. Characterisation of the immunoglobulin variable region gene usage encoding the murine anti-ganglioside antibody repertoire. *J. Neuroimmunol.* **165**, 92–103 (2005).
29. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
30. McCoy, A.J. et al. *Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
31. Collaborative Computational Project. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
32. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
33. Emsley, P. & Cowtan, K. *Coot*: model building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
34. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

Drug Design

A Potent Trivalent Sialic Acid Inhibitor of Adenovirus Type 37 Infection of Human Corneal Cells**

*Sara Spjut, Weixing Qian, Johannes Bauer, Rickard Storm, Lars Frängsmyr, Thilo Stehle, Niklas Arnberg, and Mikael Elofsson**

Viruses of the *Adenoviridae* family are widespread in society and are associated with a wide variety of clinical symptoms in humans, including respiratory, gastrointestinal, and ocular diseases.[1] Epidemic keratoconjunctivitis (EKC) is a severe ocular infection and is caused by the highly contagious adenoviruses Ad8, Ad19, and Ad37.[1] Besides keratitis and conjunctivitis, other common symptoms of EKC are pain, lacrimation, red and swollen eyes, as well as decreased vision that may last for months or even years.[1] No antiviral drugs are currently available for the treatment of EKC or any other infection caused by adenoviruses. The initial event leading to EKC is binding of the viruses to glycans that contain sialic acid moieties on epithelial cells in the cornea or conjunctiva through trimeric fiber structures extending from the viral particles.[2,3] The receptor-binding domain, the fiber knob, is located at the C terminus of each fiber and contains three separate pockets that each can accommodate one sialic acid residue. Ad37 was recently shown to bind to cell-surface glycoproteins carrying a glycan structure similar to the GD1a ganglioside.[4] The GD1a glycan is a branched hexasaccharide

[*] Dr. S. Spjut, Dr. W. Qian, Prof. M. Elofsson
Department of Chemistry
Umeå Centre for Microbial Research (UCMR) and
Laboratory for Molecular Infection Medicine Sweden (MIMS)
Umeå University, 90187 Umeå (Sweden)
E-mail: mikael.elofsson@chem.umu.se

J. Bauer, Prof. T. Stehle
Interfaculty Institute of Biochemistry, University of Tübingen
72076 Tübingen (Germany)

R. Storm, Dr. L. Frängsmyr, Prof. N. Arnberg
Department of Clinical Microbiology, Division of Virology
MIMS, Umeå University, 90187 Umeå (Sweden)

Prof. T. Stehle
Department of Pediatrics
Vanderbilt University School of Medicine
Nashville 37232, TN (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201101559.

with a terminal sialic acid residue on each of its two arms. Structural studies showed that the two sialic acid moieties dock into two of three sialic acid binding sites in the trimeric knob of the Ad37 fiber protein. Most likely, multiple fiber proteins simultaneously engage several host-cell epitopes containing terminal sialic acids; internalization and subsequent infection follow. If these sialic acid–protein interactions can be blocked, for example, by a multivalent sialic acid conjugate, infection might be prevented.

To date, several carbohydrate-based or glycomimetic drugs have reached the market; however, the development of additional therapies is still hampered by challenges such as poor absorption and/or rapid elimination.[5] The topical administration of sialic acid conjugates directly to the eye, that is, the site of infection, circumvents many of the pharmacokinetic hurdles and has the potential to prevent or even cure EKC. In the search for new antiviral substances against Ad37, we synthesized and evaluated multivalent human serum albumin (HSA) conjugates of both 3′-sialyllactose and sialic acid as adenoviral inhibitors.[6,7] These conjugates efficiently inhibited Ad37 cell attachment and the subsequent infection of human corneal epithelial (HCE) cells. Both types of conjugates were equally efficient as Ad37 inhibitors.[6,7] From the crystal structure of the fiber-knob protein as a complex with sialyllactose, it was evident that the sialic acid acetamide group is positioned in a relatively large hydrophobic pocket.[8] To improve the potency of the more advantageous multivalent sialic acid conjugates, we used structure-based design and synthesized a library of ten *N*-acyl-modified sialic acid derivatives with the overall goal of improving hydrophobic interactions and thus affinity and efficacy.[9] Unfortunately, none of the designed conjugates were as potent as the original sialic acid–HSA conjugate, although X-ray crystallography revealed that the modified saccharides interacted with the fiber-knob protein as expected.[9]

On the basis of the structural features of the interaction of the GD1a glycan with the Ad37 knob,[4] and with our previous results[6,7,9] in mind, we then designed and synthesized sialic acid containing compounds by using small non-protein scaffolds. The crystal structure of the fiber-knob protein shows that the three known sialic acid binding sites are separated by distances of about 10 Å. We therefore considered the design of a compact and rigid scaffold decorated with three correctly positioned sialic acids as too complex. Instead we selected the three small and flexible scaffolds tris(2-aminoethyl)amine (**2**), 2-(aminomethyl)-2-methyl-1,3-propanediamine (**3**), and 2,2-diaminomethyl-1,3-propanediamine (**4**) for conjugation with the sialic acid derivative **1**[7] through

Scheme 1. Synthesis of tri- and tetravalent sialic acid compounds. The sialic acid squaric decyl ester **1** was coupled to the scaffolds **2**, **3**, and **4** to form the compounds **ME0322**, **ME0323** and **ME0324**: a) *N*,*N*-diisopropylethylamine, MeOH, room temperature, 2–8 days, 30–69%.

the use of squaric acid chemistry and thus prepared **ME0322**, **ME0323**, and **ME0324** in modest to good yields (Scheme 1; see also the Supporting Information). We reasoned that longer flexible spacers would enable all three binding pockets to be occupied simultaneously by one molecule.

The effects of **ME0322**, **ME323**, **ME0324**, sialic acid, and a 17-valent sialic acid–HSA conjugate were first evaluated in
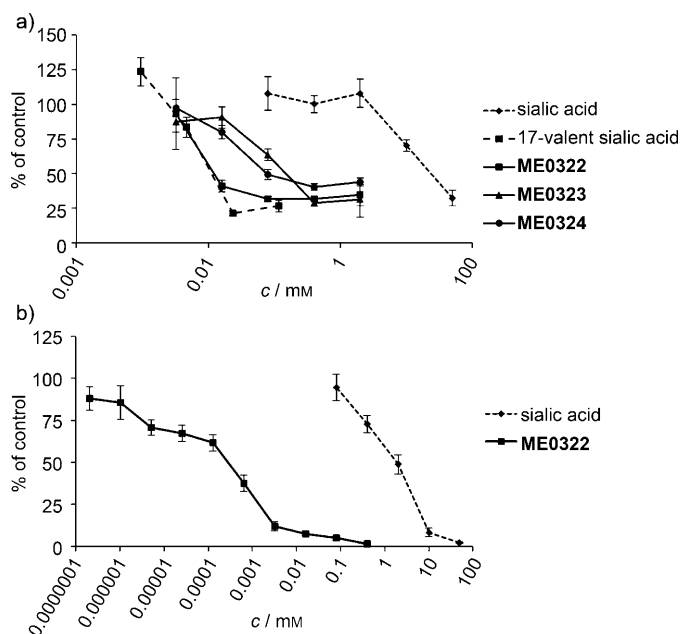
virus-binding experiments (see the Supporting Information).[6,7,9] **ME0322**, **ME0323**, and **ME0324** all inhibited the attachment of Ad37 virions to HCE cells in a dose-dependent manner and were at least two orders of magnitude more effective than sialic acid. Importantly, the most potent compound, **ME0322**, was as efficient as the 17-valent sialic acid–HSA conjugate (Figure 1a).[7] To firmly establish the potential of **ME0322** as an anti-adenoviral agent, we evaluated the compound further in a cell-based infection assay (see the Supporting Information).[6,7,9] Compound **ME0322** proved to be very potent and inhibited the infection of HCE cells by Ad37 virions with an IC_{50} value of 0.38 μM (Figure 1b). Remarkably, the trivalent compound **ME0322** was approximately four orders of magnitude more potent than sialic acid (Figure 1b) and substantially more potent than 3′-sialyllactose–HSA and sialic acid–HSA conjugates.[6,9]

To determine the structural features of the Ad37 fiber knob–**ME0322** interaction, we solved the crystal structure of this complex at a resolution of 2.4 Å (Figure 2; see also the Supporting Information, including Table S1). Only the terminal sialic acid residues of **ME0322** are visible in the final electron-density map (see Figure S1 in the Supporting Information). We did not observe electron density for the rest of the compound, probably because the flexible linkers do not make defined contacts with the protein. It is, however, likely that all three sialic acids in a fiber knob belong to the same trivalent compound. Analysis of the crystal packing shows that the distances between sialic acid binding sites in different knobs are too large to be bridged by a single **ME0322** molecule (data not shown). The binding of all sialic acid residues and their interactions with their respective protein chains are identical to the previously established binding mode between the sialic acid residues of sialyl-α(2,3)-lactose or the GD1a hexasaccharide and the Ad37 fiber knob.[4,8]

Figure 1. Effect of sialic acid containing compounds on adenovirus type 37 binding to and infection of human corneal cells. a) Extent of virion binding in the presence of inhibitors at different concentrations. b) Extent of infection at different concentrations of the inhibitors sialic acid and **ME0322**. The control is the value obtained for binding or infection in the absence of an inhibitor.

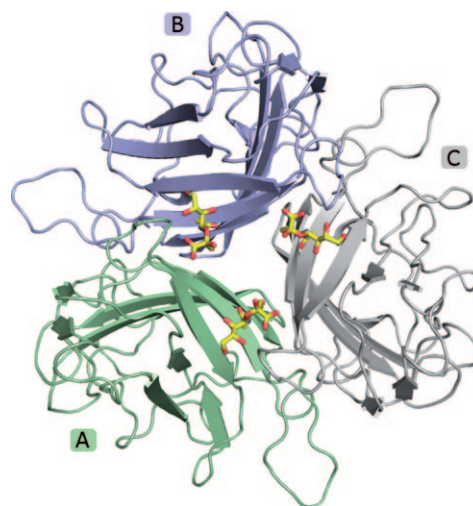**Figure 2.** Overall structure of the Ad37 fiber head as a complex with **ME0322**. The three Ad37 chains A, B, and C are shown as ribbon tracings and colored green, blue, and gray, respectively. The terminal sialic acid moietes of **ME0322** are bound on top of the fiber head and are shown in stick representation, with carbon atoms in yellow, oxygen atoms in red, and nitrogen atoms in blue. The spacers and scaffold are not shown.

Finally, we investigated the binding affinity of the trivalent sialic acid conjugate **ME0322** for Ad37 fiber knobs immobilized on a CM5 sensor chip by surface plasmon resonance (see the Supporting Information, including Figure S2). The binding was shown to fit to a simple one-to-one binding model, and the calculated K_d value for the interaction of the Ad37 fiber-knob protein with **ME0322** was calculated to be 14 μm. The interaction of the GD1a hexasaccharide with the Ad37 fiber-knob protein, on the other hand, follows a two-to-one binding model with $K_d = 19$ and 265 μm.[4] Our results suggest that the high-affinity interaction results from the occupation of two of the three sialic acid binding pockets by the two terminal sialic acids of one GD1a hexasaccharide, and that another, unknown site is engaged in the low-affinity interaction. **ME0322** and the GD1a hexasaccharide thus bind to the fiber-knob protein with similar affinities. Interestingly, the GD1a hexasaccharide is a poor inhibitor of cell infection ($IC_{50} = 0.7$ mm),[4] in contrast to **ME0322**, which is approximately four orders of magnitude more potent. The number of components and processes in the cell-based assay, however, makes it difficult to directly relate calculated affinities obtained by surface plasmon resonance analysis to potency in the infection assay.

In conclusion, we have synthesized tri- and tetravalent sialic acid compounds and evaluated them in an Ad37 cell-binding assay. The most promising trivalent compound, **ME0322**, was subsequently shown to be a very potent inhibitor of Ad37 infection of human ocular cells. Our functional and structural data show that **ME0322** efficiently blocks the adenovirus cell-binding protein, the fiber knob.

Therefore, such compounds offer promise as antiviral drugs for the topical treatment of EKC.

[1] W. S. M. Wold, M. S. Horwitz in *Fields Virology, Vol. 2* (Eds.: D. M. Knipe, P. M. Howley), Lippincott Williams & Wilkins, Philadelphia, **2007**, pp. 2395–2396.
[2] N. Arnberg, K. Edlund, A. H. Kidd, G. Wadell, *J. Virol.* **2000**, *74*, 42–48.
[3] N. Arnberg, A. H. Kidd, K. Edlund, F. Olfat, G. Wadell, *J. Virol.* **2000**, *74*, 7691–7693.
[4] E. C. Nilsson, R. J. Storm, J. Bauer, S. M. C. Johansson, A. Lookene, J. Ångstrom, M. Hedenström, T. L. Eriksson, L. Frängsmyr, S. Rinaldi, H. J. Willison, F. P. Domellöf, T. Stehle, N. Arnberg, *Nat. Med.* **2011**, *17*, 105–109.
[5] B. Ernst, J. L. Magnani, *Nat. Rev. Drug Discovery* **2009**, *8*, 661–677.
[6] S. M. C. Johansson, N. Arnberg, M. Elofsson, G. Wadell, J. Kihlberg, *ChemBioChem* **2005**, *6*, 358–364.
[7] S. M. C. Johansson, E. C. Nilsson, M. Elofsson, N. Ahlskog, J. Kihlberg, N. Arnberg, *Antiviral Res.* **2007**, *73*, 92–100.
[8] W. P. Burmeister, D. Guilligay, S. Cusack, G. Wadell, N. Arnberg, *J. Virol.* **2004**, *78*, 7727–7736.
[9] S. Johansson, E. Nilsson, W. X. Qian, D. Guilligay, T. Crepin, S. Cusack, N. Arnberg, M. Elofsson, *J. Med. Chem.* **2009**, *52*, 3666–3678.

Viruses and sialic acids: rules of engagement

Ursula Neu[1], Johannes Bauer[1] and Thilo Stehle[1,2]

Viral infections are initiated by specific attachment of a virus particle to receptors at the surface of the host cell. For many viruses, these receptors are glycans that are linked to either a protein or a lipid. Glycans terminating in sialic acid and its derivatives serve as receptors for a large number of viruses, including several human pathogens. In combination with glycan array analyses, structural analyses of complexes of viruses with sialylated oligosaccharides have provided insights into the parameters that underlie each interaction. Here, we compare the currently available structural data on viral attachment proteins in complex with sialic acid and its variants. The objective is to define common parameters of recognition and to provide a platform for understanding the determinants of specificity. This information could be of use for the prediction of the location of sialic acid binding sites in viruses for which structural information is still lacking. An improved understanding of the principles that govern the recognition of sialic acid and sialylated oligosaccharides would also advance efforts to develop efficient antiviral agents.

Addresses
[1] Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler-Str. 4, D-72076 Tübingen, Germany
[2] Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN 37232, United States

Corresponding author: Stehle, Thilo (thilo.stehle@uni-tuebingen.de)

Introduction

The monosaccharide sialic acid decorates all eukaryotic cell surfaces, capping many different oligosaccharide structures on N-linked and O-linked glycoproteins as well as on glycolipids [1]. Glycans terminating in sialic acid have emerged as a key class of receptors for an impressive number of viruses, many of which are human pathogens. The highly pathogenic Influenza A, B and C viruses as well as the human parainfluenza viruses attach to sialic acids (reviewed in [2–4]). Coxsackievirus A24 variant and enterovirus 70, which cause Acute Hemorrhagic Conjunctivitis and have pandemic potential, also attach to sialylated oligosaccharides [5,6]. Epidemic Keratoconjunctivitis (EKC) has been linked to several human D-type adenoviruses, and one of

these has recently been shown to attach to the disialylated GD1a motif [7••]. The human JC and BK polyomaviruses (JCV and BKV, respectively) cause a fatal demyelinating disease and kidney graft loss, respectively, in immunocompromised individuals. Both viruses use glycans terminating in sialic acid as their receptors [8••,9]. Moreover, the recently identified Merkel cell polyomavirus, a human oncovirus, likely uses the trisialylated ganglioside GT1b as a receptor [10•,11], and other mammalian polyomaviruses such as Simian Virus 40 (SV40) and murine polyomavirus (Polyoma) also bind glycans terminating in sialic acid [12,13•,14]. Most human noroviruses, the causative agents of violent gastrointestinal illnesses, attach to non-sialylated histo-blood group antigens, in contrast to murine norovirus, which binds to a sialylated oligosaccharide [15]. However, some strains of human noroviruses were recently shown to bind to the sialyl-Lewis X motif as well [16•]. Rotaviruses cause severe gastroenteritis in children. They have long been classified into strains that can be inhibited by neuraminidase treatment, which cleaves sialic acid from glycan sequences on host cells, and those that are insensitive to it [17,18]. Neuraminidase-insensitive strains were presumed to use non-sialylated receptors. Interestingly, the 'neuraminidase-insensitive' rotavirus strain Wa was recently shown to attach to the ganglioside GM1, which carries a branching sialic acid [19••]. Because of its branched structure, this particular carbohydrate is difficult to cleave with neuraminidases [20,21]. SV40 had likewise been presumed to attach to a non-sialylated carbohydrate [22–24] before GM1 was identified as its receptor [12]. Sialic acid, therefore, has to be considered as a possible receptor component even for viruses whose infectivity cannot be modulated by treatment of cells with commonly used neuraminidases.

The structure of the most common sialic acid in humans, α-5-N-acetyl-neuraminic acid (Neu5Ac), features four protruding functional groups (carboxylate, hydroxyl, N-acetyl and glycerol functions). Compared to more simple monosaccharides, the large number of functional groups enables sialic acids to participate in an unparalleled number of hydrogen bonds, salt bridges and non-polar interactions. Since sialic acid is typically located at the terminus of a glycan, its functions are easily accessible for interactions. Perhaps it is not surprising, therefore, that the sialic acid itself serves as the major point of contact with the glycan-binding viral attachment protein in all cases where structural information of sufficient resolution is available [8••,13•,25–34,35•,36•,37–40,41•].

In this review, we compare binding modes of sialic acids and sialylated receptors in representative structures of virus-receptor complexes in order to derive parameters

that guide the recognition of sialic acid and its derivatives. Such parameters could be useful for the prediction of new sialic acid binding sites in viral proteins, or of altered modes of sialic acid (and its derivatives) binding in different viral serotypes and strains.

Attachment to terminal sialic acid

Several structures of viral attachment proteins in complex with sialylated compounds have been determined recently, providing new insights into viral specificity for glycan receptors [8,13•,35•,36•,37–39,41•]. Taken together, the known structures now form a large database that is suitable for the closer examination of contacts in order to compare the modes of interaction and attempt to define common principles of sialic acid recognition. We have investigated here the mode of Neu5Ac binding for ten different viral attachment proteins (Figure 1): the

hemagglutinins (HAs) of Influenza A and B viruses [26,34], the adenovirus serotype 37 (Ad37) fiber knob [7••], the canine adenovirus serotype 2 (cAd2) fiber knob [39], the major capsid proteins VP1 of the polyomaviruses Polyoma [29], SV40 [13•], and JCV [8], the attachment protein σ1 of human type 3 orthoreovirus [41•], the attachment protein VP8* of Rhesus Rotavirus [30], and the hemagglutinin-neuraminidase (HN) of Newcastle Disease Virus (NDV) [31]. As the interactions with terminal sialic acid are very similar among different types of Influenza A HAs, the structure of the H3 type [26] was chosen to represent this group. In addition, we have also analyzed contacts in the hemagglutinin-esterase-fusion (HEF) protein of Influenza C virus [27] and the hemagglutinin-esterase (HE) proteins of Bovine Coronavirus (BCoV) [35], Bovine Torovirus (BToV) [36•] and Porcine Torovirus (PToV) [36•] (Figure 2). These four proteins

Figure 1
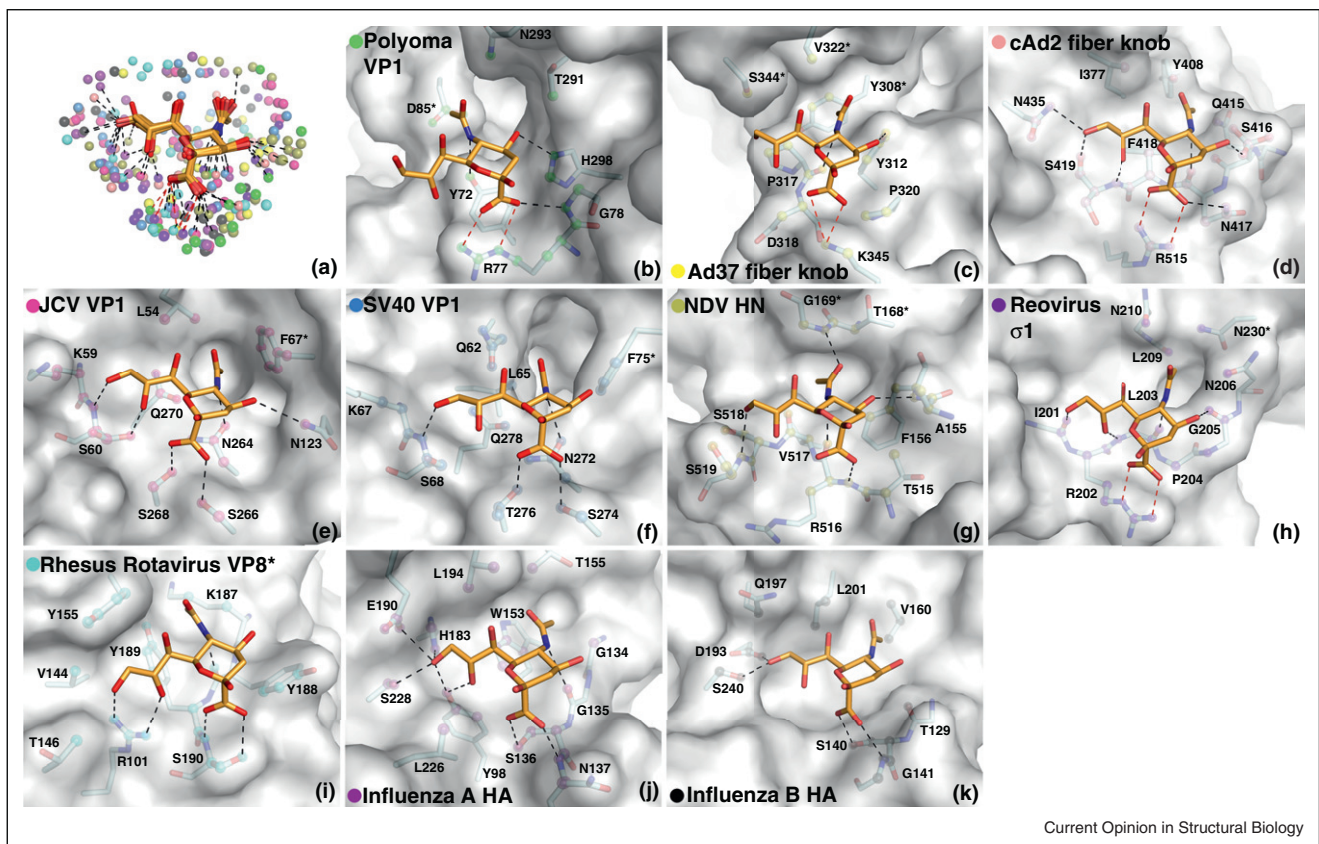
Interactions of viral attachment proteins with terminal Neu5Ac. **(a)** Contacts of viral proteins with terminal Neu5Ac. The complex structures shown in panels **(b–k)** were superposed using the terminal sialic acid residues. The Neu5Ac residues are shown in stick representation, with carbons colored orange, oxygens red, and nitrogens blue. All protein atoms within a 4.0 Å radius of Neu5Ac atoms are displayed as spheres and colored according to the color code depicted in panels (b–k). Hydrogen bonds and salt bridges are represented with black and red dashed lines, respectively. **(b–k)** Neu5Ac binding sites of Polyoma VP1 (b, pdb 1VPS [29]), Ad37 fiber knob (c, pdb 3N0I [7]), cAd2 fiber knob (d, pdb 2WBV [39]), JCV VP1 (e, pdb 3NXD [8]), SV40 VP1 (f, pdb 3BWR [13]), NDV HN (g, pdb 1USR [31]), Reovirus σ1 (h, pdb 3S6X [41]), Rhesus Rotavirus VP8* (i, pdb 1KQR [30]), Influenza A HA (j, 1HGG [26]), and Influenza B HA (k, pdb 2RFT [34]). In all cases, Neu5Ac is shown in stick representation and colored as in panel (a). The protein surfaces are colored gray, with residues interacting with Neu5Ac shown in stick representation and colored by element. Protein atoms within a 4.0 Å radius around Neu5Ac are highlighted with colored spheres. In cases where the Neu5Ac binding site is formed by two protein chains, one of the chains is denoted with an asterisk.

Figure 2
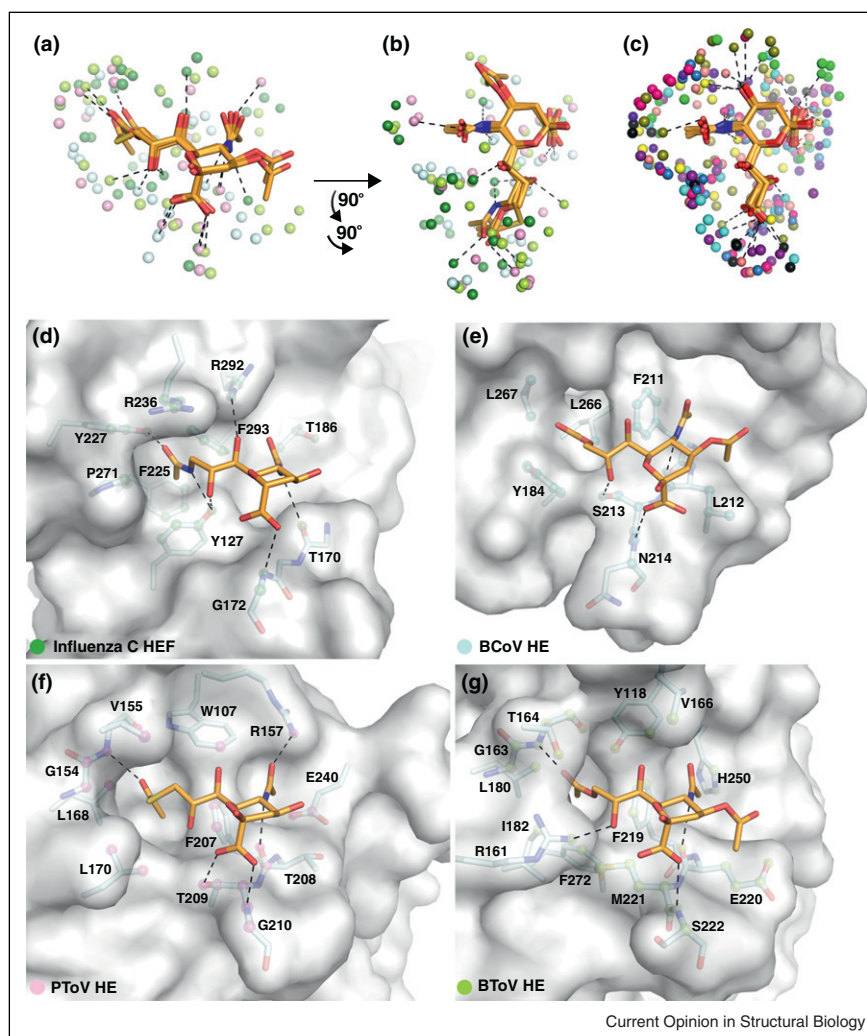
Interactions of viral attachment proteins with *O*-acetylated Neu5Ac. **(a)** Contacts of viral proteins with terminal *O*-acetylated Neu5Ac. Complex structures were superposed on the terminal substituted Neu5Ac. Interacting protein atoms as well as hydrogen bonds and salt bridges are shown as in Figure 1a. Top view of *O*-acetylated **(b)** and unsubstituted **(c)** Neu5Ac, with contacting protein atoms shown as colored spheres. The views are rotated from those in Figure 1 and (a) by two 90° rotations, one around a horizontal and one around a vertical axis. **(d–g)** Binding sites for *O*-acetylated Neu5Ac in Influenza C HEF (d) [27], BCoV HE (e, pdb 3CL5 [35]), PToV HE (f, pdb 3I27 [36]) and BToV HE (g, pdb 3I1L [36]). The structures are displayed as in Figure 1b–k.

bind to derivatives of Neu5Ac that are *O*-acetylated at position 9 or at positions 4 and 9. In cases where the attachment protein also has receptor-destroying enzymatic activity, such as in the HN and HE(F) proteins, only the sites that can clearly be attributed to attachment are considered here, thus excluding the dual function neuraminidase site of some HN proteins [32,42,43].

The investigated viral attachment proteins are, for the most part, not homologous to one another and belong to unrelated viruses that differ in envelope structure and genome type. Nevertheless, their interactions with sialic acid display striking similarities (Figures 1 and 2). In all complexes, the sialic acid adopts essentially the same

conformation, namely a *trans* conformation of the 5-*N*-acetyl group and an α-conformation at the anomeric carbon, which is dominant in biological oligosaccharides. Interestingly, all attachment proteins, including the ones that bind to *O*-acetylated compounds, make extensive contacts with one face of the sialic acid ring, while the other face is engaged by only few contacts (Figures 1a and 2a–c). A likely reason for this preference is the formation of two key contacts that are formed in all complexes in a similar manner. One of these contacts involves the negatively charged carboxylate group, which is most often recognized by two parallel hydrogen bonds or a salt bridge. Each of the analyzed proteins donates at least one hydrogen bond to a carboxylate oxygen atom. The

second contact involves the nitrogen atom in the *N*-acetyl group. With only one exception, all proteins receive a hydrogen bond from this nitrogen atom. The spatial arrangement of carboxyl group and the *N*-acetyl nitrogen thus helps distinguish sialic acids from other monosaccharides. Both groups project from the same face of the sialic acid ring, accounting for the preferential binding of this face of the monosaccharide. Apart from these two key interactions, the proteins engage in different hydrogen bonding patterns to various hydroxyl groups of the glycerol chain or the ring, or to additional acetyl substituents.

Examination of van der Waals interactions between the viral attachment proteins and Neu5Ac reveals that only about 50% of the contact surface of Neu5Ac participates in such contacts (Figures 1a and 2c). The resulting shape resembles a rimmed imprint of the binding face of Neu5Ac on the protein surface (Figures 1a and 2c). In all complexes, van der Waals interactions are formed with the methyl group of the *N*-acetyl chain. However, the different proteins interacting with Neu5Ac sample different epitopes on the Neu5Ac contact surface. For example, JCV and SV40 VP1, Rhesus Rotavirus VP8*, and Influenza A HA all center their van der Waals contacts on the glycerol and *N*-acetyl chains (Figure 1e,f,i,j). The surfaces of all four viruses feature subtle protrusions that separate the recessed areas in which the glycerol and *N*-acetyl chains are bound. Polyoma VP1, on the other hand, mainly contacts Neu5Ac from the other side, and does not interact with the glycerol chain at all (Figure 1b). Examination of the binding surfaces demonstrates that shape complementarity is an important factor in the engagement of sialic acid. As the contact areas are quite small and the sialic acids are partially exposed to solvent, adding or removing a single contact can thus have significant effects on the affinity of a given virus for sialic acid or its variants.

Recognition of sialic acid variants

The parent compound of Neu5Ac, neuraminic acid, can feature numerous modifications that give rise to over 40 different known sialic acid variants [44,45]. Several of these modifications are predominantly found on specific cell types and tissues, or in selected species. It is perhaps not surprising, therefore, that some viruses exploit this divergence and preferentially recognize sialic acids other than Neu5Ac. The database of viral protein structures contains few examples of viruses attaching to *O*-acetylated Neu5Ac [27,35,36•], but their analysis is nevertheless informative (Figure 2). While the key hydrogen bonds to one face of the sialic acid are the same as described for Neu5Ac, the distribution of van der Waals contacts is somewhat altered. In the four complexes, the majority of van der Waals contacts are centered around the unique 9-*O*-acetyl groups as well as the adjacent side of the *N*-acetyl group, while the opposite side of the ring does not engage in as many interactions (Figure 2b). The

9-*O*-acetyl group inserts deeply into tight-fitting protein cavities, providing selectivity for sialic acids modified in this manner. Recognition of different sialic acids is also a likely cause of changes in tropism and host range. The interactions of SV40 with GM1 ganglioside containing α-*N*-5-glycolyl neuraminic acid (Neu5Gc), a sialic acid present in simians but not humans, illustrate this point [46]. SV40 VP1 features a large pocket near the Neu5Ac *N*-acetyl group (Figures 1f and 3b), and it is tempting to speculate that this pocket serves to accommodate the additional hydroxyl group of Neu5Gc [13•]. VP1 of the human JCV (Figures 1e and 3a), whose sialic acid binding site is largely similar to that of SV40 VP1, features a much smaller pocket that likely prefers the smaller human Neu5Ac over the simian Neu5Gc.

Conservation of sialic acid binding sites

Sialic acid binding sites are often highly conserved in homologous viruses. This is evident when comparing different HA types of Influenza A, the capsid proteins of JCV and SV40 (Figure 3a,b), or the HE proteins of PToV and BToV (Figure 2f,g). In all three cases, the sialic acid engages the two homologous proteins using similar contacts and is therefore bound in the same orientation and position. However, at least one example exists where homologous proteins, the Ad37 and cAd2 fiber knobs, bind sialic acid at different locations (Figure 1c,d). Interestingly, there are several examples of highly homologous proteins that bind sialic acid at the same site but in different orientations. The VP1 proteins of Polyoma and SV40, for example, feature a very high level of sequence identity, and they bind sialic acid in generally similar areas on the protein surface. However, the orientations of the bound sialic acids differ markedly (Figure 1b,f) [13•]. Similarly, Influenza C HEF and BCoV HE bind sialic acid at the same position, but again in different orientations with respect to the proteins and with contacts provided by different structural elements (Figure 2d,e) [35]. These two examples demonstrate the need for caution when modeling interactions with sialic acid based on a homologous structure.

Specificity for sialic acid in different contexts

As sialic acid is ubiquitous at the cell surface, interactions with subsequent carbohydrates are typically employed to define specificity and tropism. Glycan microarrays have been highly useful in revealing the determinants of such interactions [47]. The critical role of the context of the sialic acid — linkage type, as well as length, sequence and conformational preferences of the remaining oligosaccharide chain — is perhaps best illustrated by its influence on the host range of influenza viruses. Briefly, human Influenza A viruses engage long glycans terminating in α2,6-linked sialic acid that preferentially adopt a bent conformation and that are expressed extensively in the upper airway epithelia of humans. Avian strains predominantly recognize shorter glycans that adopt a linear

Figure 3
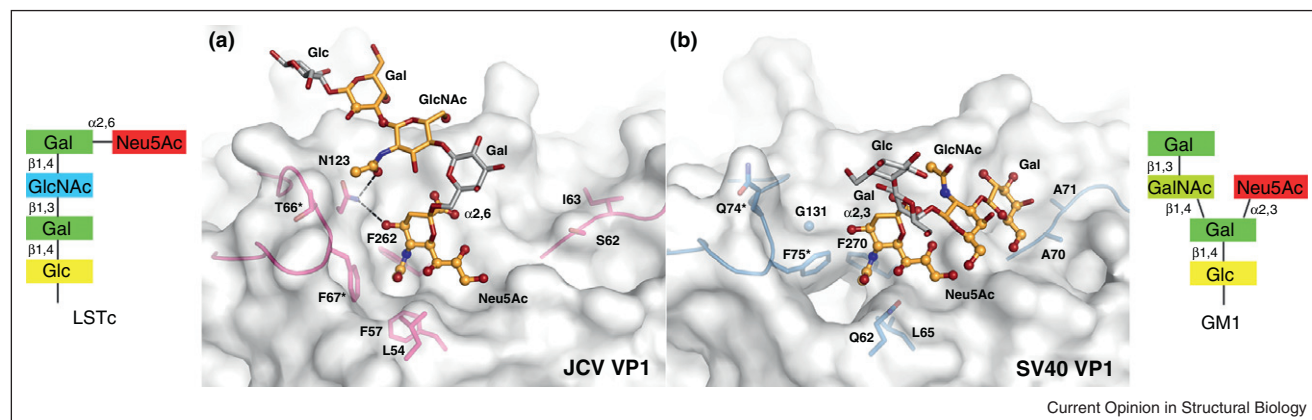
Different sialic acid and context specificity of SV40 and JCV VP1. **(a)** JCV VP1 in complex with LSTc oligosaccharide. **(b)** SV40 VP1 in complex with GM1 oligosaccharide. The oligosaccharides are shown in stick representation and colored by element, with oxygens in red and nitrogens in blue. Monosaccharides that approach the protein closer than 4.0 Å are colored in bright orange, while those not contacting the protein are colored gray. Oligosaccharide atoms within a 4.0 Å radius around the proteins are highlighted as spheres. The protein surface is shown in gray. Residues that define the different oligosaccharide specificities of the two proteins are shown as sticks and colored blue and pink for SV40 and JCV VP1, respectively. Residues from a different polypeptide chain are denoted with an asterisk.

conformation and that often contain α2,3-linked linked sialic acid [48•]. The vast database on Influenza A HA structures in complex with sialylated ligands, and concurrent glycan array analyses, has been the subject of several excellent recent reviews [2,3], and will therefore not be discussed in detail here. However, glycan array screening has recently helped to unravel the identities of sialylated glycan receptors for two pathogenic human viruses, and structural biology has defined the nature of interaction in both cases [8]. We review below each of these two examples, which illuminate the importance of the context in which the terminal sialic acid is placed.

## Achieving specificity through a limited set of additional interactions

Several members of the polyomavirus family use sialylated receptors for cell attachment. Crystal structures of two members of the family in complex with their cognate receptors have been determined recently: SV40 VP1 has been crystallized in complex with the oligosaccharide portion of its ganglioside receptor GM1 [13•], whereas the structure of the VP1 protein of human JCV has been solved with the pentasaccharide receptor fragment LSTc (Lacto-series tetrasaccharide c) [8]. Both receptors feature terminal Neu5Ac, which is α2,3-linked in the branched GM1 molecule and α2,6-linked in the linear LSTc structure (Figure 3). In each case, glycan array screening has unequivocally identified the type of the receptor [8,13•]. Moreover, although both GM1 and LSTc were present on the arrays, JCV VP1 failed to interact with GM1, and SV40 VP1 also did not recognize the LSTc compound. Thus, both proteins are highly specific for their cognate receptors. A comparison of

the two structures shows that the sialic acid portions of the two receptors form largely equivalent interactions with their respective proteins (Figures 1e,f and 3). The remarkable specificity for each receptor can be attributed to contacts that involve the remaining parts of the oligosaccharides. The LSTc compound assumes a bent conformation, forming additional contacts via the *N*-acetyl group of its third sugar, GlcNAc, to N123 of JCV (Figure 3a). An α2,3-linked Neu5Ac would not adopt a similarly bent conformation, explaining why sialylparagloboside, which is identical to LSTc except for its α2,3-linked Neu5Ac, does not bind JCV. Modeling LSTc into the SV40 VP1 binding site by superimposing the two sialic acid structures suggests that LSTc could be tolerated by SV40. However, as the residue equivalent to N123 in JCV is a glycine (G131) in SV40, no favorable contacts to the GlcNAc residue can be generated (Figure 3b). The inability to form such an interaction most likely explains why SV40 cannot bind LSTc. It, therefore, appears that the formation of a very small number of contacts is largely responsible for defining the specificity of VP1 for LSTc. The reverse combination, a GM1 ligand bound to JCV, would likely be disfavored due to steric clashes with JCV VP1 residue S62, which is an alanine in SV40.

Achieving specificity through multivalent binding

Because of their small contact surfaces and solvent-exposed binding sites, interactions between individual viral attachment proteins and sialylated oligosaccharides are typically of low affinity, with dissociation constants in the millimolar range [13•,28,49,50]. In many cases,

Figure 4
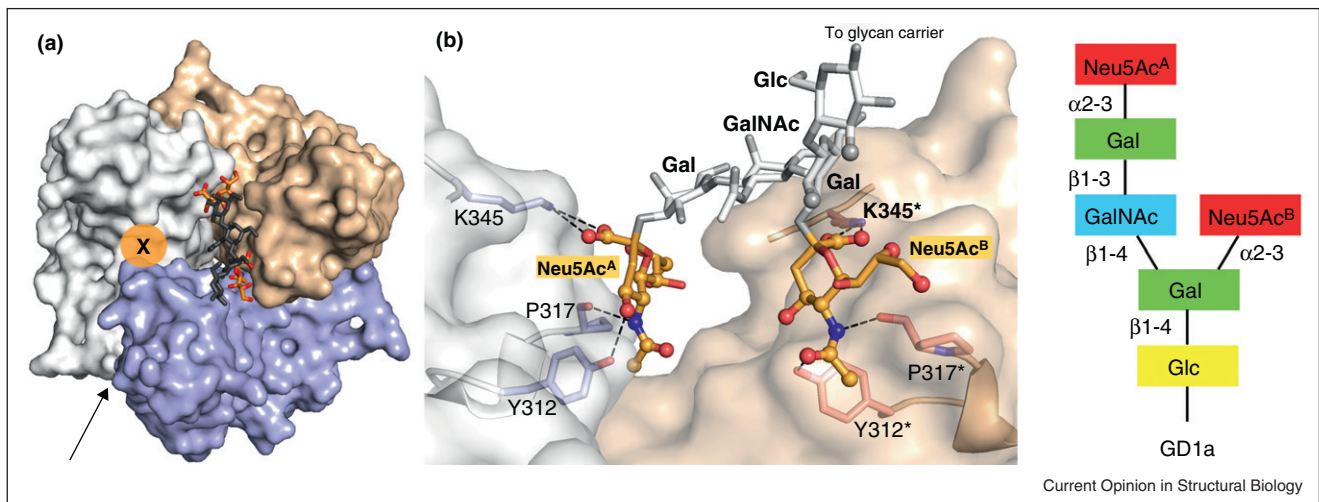
Binding of GD1a glycan to the Ad37 fiber knob protein. (a) Structure of Ad37 fiber knob in complex with the GD1a oligosaccharide [7]. The three different Ad37 knob protomers are shown in surface representation and are colored gray, light red and blue. The GD1a glycan is drawn in stick representation, with both Neu5Ac residues highlighted in color (carbons in orange, oxygens in red and nitrogens in blue). The bridging glycan residues are shown in dark gray. The third binding site (marked by 'X') is blocked due to crystal contacts. The arrow indicates the viewing direction shown in panel (b). (b) Interactions between Ad37 fiber knob residues and GD1a. Two different Ad37 fiber knob protomers are shown in transparent surface representation (white and light red). The third protomer is not shown for clarity. Ad37 residues contacting GD1a are shown in stick representation, with oxygens in red and nitrogens in blue. The GD1a glycan is shown in stick representation, with both terminal Neu5Ac residues highlighted in color (carbons in orange, oxygens in red and nitrogens in blue). The bridging glycan residues are shown in gray. Glycan atoms within a distance of 4 Å to Ad37 protein atoms are drawn as spheres. Hydrogen bonds to GD1a glycan are represented with black dashes. Residues from different protomers are denoted with an asterisk.

high-affinity adherence to the target cell is achieved through the utilization of several low-affinity binding sites. However, receptor clustering is not always necessary to achieve higher-affinity binding. The interaction of Ad37 with its recently identified glycan receptor GD1a illustrates another strategy. It has long been established that Ad37 fiber knobs bind receptors terminating in sialic acid [28,51], but the nature of the glycan has remained elusive. Glycan array screening has recently revealed that Ad37 fiber knobs specifically recognize the oligosaccharide GD1a, a disialylated compound that features two branches, each terminating in sialic acid [7]. A structural analysis of the trimeric Ad37 fiber knob in complex with GD1a established that the two terminal sialic acid residues bind to two different Ad37 fiber knob protomers in an identical manner, thus engaging two of the three possible binding sites [7] (Figure 4). This bivalent interaction results in a 250-fold higher affinity (K_d = 19 μM) [7] compared to the monovalent sialyllactose–Ad37 knob interaction (K_d = 5 mM) [28]. Thus, although each protomer in an Ad37 fiber knob would be able to bind sialic acid attached to different oligosaccharide structures, specificity for GD1a is generated by a multivalent interaction in which two protomers interact with the same receptor in an identical manner. It is conceivable that trivalent compounds that engage all three binding sites of the

Ad37 fiber knobs would have even higher affinity, thus providing a platform for the development of antiviral inhibitors. Using such a strategy, a multivalent inhibitor has been developed that is able to neutralize pentameric shiga-like toxins with very high efficiency [52]. A similar strategy could be useful to develop molecules that inhibit viral attachment proteins, which usually occur as multimers at the viral surface.

Conclusions

A large number of viruses, including many serious human pathogens, use sialylated oligosaccharides for cell attachment. Common principles of interaction can be established by comparing the sialic acid binding modes of different viruses. In most cases, interactions between a viral attachment protein and its glycan receptor involve primarily the sialic acid itself, which is bound with a relatively small contact area in a solvent-exposed region of the protein. Consistent with this, the affinities of such interactions are, at least in cases where they have been measured, very low. Nevertheless, many of the viruses discussed here achieve remarkable specificity for a single type of sialylated oligosaccharide by establishing a small number of auxiliary interactions with functional groups that lie beyond the sialic acid, and by excluding some possible ligands through steric clashes. The auxiliary interactions generally involve fewer hydrogen bonds

and bury a smaller amount of surface compared to the interactions that involve the sialic acid itself. It thus appears that many viruses use the unique properties of sialic acid as a 'hook' that allows them to adhere to the cell, and modulate binding in different strains or families by subtly altering structural elements in the vicinity of this hook. In a (so far) unique variation of this strategy, the Ad37 knob establishes selectivity for its GD1a glycan receptor by multivalent binding to a single receptor carrying two terminal sialic acid moieties, thus adhering to two identical 'hooks' separated by a defined spacer. The prominence of sialic acid in viral attachment may form a basis for new approaches to combat viruses. Compounds that mimic sialic acid have already proved useful as inhibitors of the influenza virus neuraminidase [53] and can also efficiently inhibit the receptor-binding site of the Influenza A virus hemagglutinin [54]. The structural analysis of the Ad37–GD1a interaction has also led to the design and synthesis of a trivalent compound designed to block attachment of adenoviruses that cause EKC [55]. Glycan microarrays have been extraordinarily useful in identifying the correct receptors for many viral proteins [8,13•,47,56••], which is a prerequisite for structural studies. However, proper interpretation of the information provided by glycan array screening and structural analyses requires affinity data. Such data are often difficult to obtain and compare, and they are currently lacking for many complexes. Being able to correlate affinity measurements with structural data would significantly advance the design of antiviral agents, and, together with oligosaccharide expression data, help to explain viral tropism.

Acknowledgments

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
•• of outstanding interest

1. Varki A: **Sialic acids in human health and disease**. *Trends Mol Med* 2008, **14**:351-360.

2. Gamblin SJ, Skehel JJ: **Influenza hemagglutinin and neuraminidase membrane glycoproteins**. *J Biol Chem* 2010, **285**:28403-28409.

3. Viswanathan K, Chandrasekaran A, Srinivasan A, Raman R, Sasisekharan V, Sasisekharan R: **Glycans as receptors for influenza pathogenesis**. *Glycoconj J* 2010, **27**:561-570.

4. Villar E, Barroso IM: **Role of sialic acid-containing molecules in paramyxovirus entry into the host cell: a minireview**. *Glycoconj J* 2006, **23**:5-17.

5. Nilsson EC, Jamshidi F, Johansson SM, Oberste MS, Arnberg N: **Sialic acid is a cellular receptor for coxsackievirus A24 variant, an emerging virus with pandemic potential**. *J Virol* 2008, **82**:3061-3068.

6. Alexander DA, Dimock K: **Sialic acid functions in enterovirus 70 binding and infection**. *J Virol* 2002, **76**:11265-11272.

7. Nilsson EC, Storm RJ, Bauer J, Johansson SM, Lookene A,
•• Angstrom J, Hedenstrom M, Eriksson TL, Frangsmyr L, Rinaldi S etal.: **The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis**. *Nat Med* 2011, **17**:105-109.
The authors show by glycan array screening that the receptor-recognizing knob domain of the Ad37 fiber protein specifically binds a branched hexasaccharide that is present in the GD1a ganglioside and that features two terminal sialic acids. Structural analysis reveals that the two terminal sialic acids dock into two of the three previously established sialic acid-binding sites in the trimeric Ad37 knob. The findings form a basis for the design and development of sialic acid-containing antiviral drugs for topical treatment of EKC.

8. Neu U, Maginnis MS, Palma AS, Ströh L, Feizi T, Atwood WJ,
•• Stehle T: **Structure-function analysis of the human JC polyomavirus establishes the LSTc pentasaccharide as a functional receptor motif**. *Cell Host Microbe* 2010, **8**:309-319.
Using glycan array screening, the authors identify the LSTc glycan, a sialylated pentasaccharide present on glycoproteins and glycolipids, as a specific receptor recognition motif for the human JC polyomavirus. The crystal structure of the capsid protein of this virus was solved alone and in complex with LSTc, explaining its distinct receptor specificity.

9. Low JA, Magnuson B, Tsai B, Imperiale MJ: **Identification of gangliosides GD1b and GT1b as receptors for BK virus**. *J Virol* 2006, **80**:1361-1366.

10. Feng H, Shuda M, Chang Y, Moore PS: **Clonal integration of a
• polyomavirus in human Merkel cell carcinoma**. *Science* 2008, **319**:1096-1100.
This paper describes a new human polyomavirus that is causally linked to an aggressive type of skin cancer, the Merkel cell carcinoma.

11. Erickson KD, Garcea RL, Tsai B: **Ganglioside GT1b is a putative host cell receptor for the Merkel cell polyomavirus**. *J Virol* 2009, **83**:10275-10279.

12. Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, Rapoport TA: **Gangliosides are receptors for murine polyoma virus and SV40**. *Embo J* 2003, **22**:4346-4355.

13. Neu U, Woellner K, Gauglitz G, Stehle T: **Structural basis of GM1
• ganglioside recognition by simian virus 40**. *Proc Natl Acad Sci U S A* 2008, **105**:5219-5224.
The authors use glycan array screening, affinity measurements and structural analysis to define the interactions of Simian Virus 40 with its ganglioside receptor GM1.

14. Stehle T, Yan Y, Benjamin TL, Harrison SC: **Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment**. *Nature* 1994, **369**:160-163.

15. Taube S, Perry JW, Yetming K, Patel SP, Auble H, Shu L, Nawar HF, Lee CH, Connell TD, Shayman JA et al.: **Ganglioside-linked terminal sialic acid moieties on murine macrophages function as attachment receptors for murine noroviruses**. *J Virol* 2009, **83**:4092-4101.

16. Rydell GE, Nilsson J, Rodriguez-Diaz J, Ruvoen-Clouet N,
• Svensson L, Le Pendu J, Larson G: **Human noroviruses recognize sialyl Lewis × neoglycoprotein**. *Glycobiology* 2009, **19**:309-320.
This study shows that some human norovirus strains have at least two binding specificities: one related to α1,2-fucosylated carbohydrates and another related to sialylated carbohydrates.

17. Isa P, Arias CF, Lopez S: **Role of sialic acids in rotavirus infection**. *Glycoconj J* 2006, **23**:27-37.

18. Ciarlet M, Ludert JE, Iturriza-Gomara M, Liprandi F, Gray JJ, Desselberger U, Estes MK: **Initial interaction of rotavirus strains with N-acetylneuraminic (sialic) acid residues on the cell surface correlates with VP4 genotype, not species of origin**. *J Virol* 2002, **76**:4087-4095.

19. Haselhorst T, Fleming FE, Dyason JC, Hartnell RD, Yu X,
•• Holloway G, Santegoets K, Kiefel MJ, Blanchard H, Coulson BS

et al.: **Sialic acid dependence in rotavirus host cell invasion**. *Nat Chem Biol* 2009, **5**:91-93.
Using NMR analysis and infectivity assays, the authors present strong evidence that *N*-acetylneuraminic acid is a key determinant for binding of some rotaviruses, in contrast to the widely accepted paradigm that sialic acids are irrelevant in host cell recognition by sialidase-insensitive rotaviruses.

20. Miller-Podraza H, Bradley RM, Fishman PH: **Biosynthesis and localization of gangliosides in cultured cells**. *Biochemistry* 1982, **21**:3260-3265.

21. Svennerholm L: **Chromatographic separation of human brain gangliosides**. *J Neurochem* 1963, **10**:613-623.

22. Clayson ET, Compans RW: **Characterization of simian virus 40 receptor moieties on the surfaces of Vero C1008 cells**. *J Virol* 1989, **63**:1095-1100.

23. Shah KV, Valis JD, Daniel RW: **Hemagglutination with simian papovavirus SA12**. *J Clin Microbiol* 1978, **7**:396-398.

24. Clayson ET, Compans RW: **Entry of simian virus 40 is restricted to apical surfaces of polarized epithelial cells**. *Mol Cell Biol* 1988, **8**:3391-3396.

25. Weis WI, Brown JH, Cusack S, Paulson JC, Skehel JJ, Wiley DC: **Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid**. *Nature* 1988, **333**:426-431.

26. Sauter NK, Hanson JE, Glick GD, Brown JH, Crowther RL, Park SJ, Skehel JJ, Wiley DC: **Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography**. *Biochemistry* 1992, **31**:9609-9621.

27. Rosenthal PB, Zhang X, Formanowski F, Fitz W, Wong CH, Meier-Ewert H, Skehel JJ, Wiley DC: **Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus**. *Nature* 1998, **396**:92-96.

28. Burmeister WP, Guilligay D, Cusack S, Wadell G, Arnberg N: **Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites**. *J Virol* 2004, **78**:7727-7736.

29. Stehle T, Harrison SC: **High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding**. *EMBO J* 1997, **16**:5139-5148.

30. Dormitzer PR, Sun ZY, Wagner G, Harrison SC: **The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site**. *EMBO J* 2002, **21**:885-897.

31. Zaitsev V, von Itzstein M, Groves D, Kiefel M, Takimoto T, Portner A, Taylor G: **Second sialic acid binding site in Newcastle disease virus hemagglutinin-neuraminidase: implications for fusion**. *J Virol* 2004, **78**:3733-3741.

32. Yuan P, Thompson TB, Wurzburg BA, Paterson RG, Lamb RA, Jardetzky TS: **Structural studies of the parainfluenza virus 5 hemagglutinin-neuraminidase tetramer in complex with its receptor, sialyllactose**. *Structure* 2005, **13**:803-815.

33. Blanchard H, Yu X, Coulson BS, von Itzstein M: **Insight into host cell carbohydrate-recognition by human and porcine rotavirus from crystal structures of the virion spike associated carbohydrate-binding domain (VP8*)**. *J Mol Biol* 2007, **367**:1215-1226.

34. Wang Q, Tian X, Chen X, Ma J: **Structural basis for receptor specificity of influenza B virus hemagglutinin**. *Proc Natl Acad Sci U S A* 2007, **104**:16874-16879.

35. Zeng Q, Langereis MA, van Vliet AL, Huizinga EG, de Groot RJ:
• **Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution**. *Proc Natl Acad Sci U S A* 2008, **105**:9065-9069.
A structure analysis of the coronavirus hemagglutinin-esterase in complex with its receptor establishes an evolutionary link to the influenza C-like hemagglutinin-esterase fusion protein.

36. Langereis MA, Zeng Q, Gerwig GJ, Frey B, von Itzstein M,
• Kamerling JP, de Groot RJ, Huizinga EG: **Structural basis for ligand and substrate recognition by torovirus hemagglutinin esterases**. *Proc Natl Acad Sci U S A* 2009, **106**:15897-15902.
The structure analysis of two torovirus hemagglutinin-esterases sheds light on the evolution of these proteins and provides insight into mechanisms of substrate binding, substrate recognition, and receptor selection in this important class of viral proteins.

37. Liu J, Stevens DJ, Haire LF, Walker PA, Coombs PJ, Russell RJ, Gamblin SJ, Skehel JJ: **Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957**. *Proc Natl Acad Sci U S A* 2009, **106**:17175-17180.

38. Lin T, Wang G, Li A, Zhang Q, Wu C, Zhang R, Cai Q, Song W, Yuen KY: **The hemagglutinin structure of an avian H1N1 influenza A virus**. *Virology* 2009, **392**:73-81.

39. Seiradake E, Henaff D, Wodrich H, Billet O, Perreau M, Hippert C, Mennechet F, Schoehn G, Lortat-Jacob H, Dreja H *et al.*: **The cell adhesion molecule "CAR" and sialic acid on human erythrocytes influence adenovirus in vivo biodistribution**. *PLoS Pathog* 2009, **5**:e1000277.

40. Yang H, Chen LM, Carney PJ, Donis RO, Stevens J: **Structures of receptor complexes of a North American H7N2 influenza hemagglutinin with a loop deletion in the receptor binding site**. *PLoS Pathog* 2010, **6**:e1001081.

41. Reiter DM, Frierson JM, Halvorson EE, Kobayashi T, Dermody TS,
• Stehle T: **Crystal structure of reovirus attachment protein σ1 in complex with sialylated oligosaccharides**. *PLoS Pathog* 2011, **7**:e1002166.
This is the first example of a viral protein engaging a carbohydrate receptor through a fibrous, trimeric structure formed from repeating units, the β-spiral. The ligand-binding site is formed by a compact, trimeric module that could potentially be grafted onto other trimeric proteins, endowing them with carbohydrate binding capabilities.

42. Crennell S, Takimoto T, Portner A, Taylor G: **Crystal structure of the multifunctional paramyxovirus hemagglutinin-neuraminidase**. *Nat Struct Biol* 2000, **7**:1068-1074.

43. Lawrence MC, Borg NA, Streltsov VA, Pilling PA, Epa VC, Varghese JN, McKimm-Breschkin JL, Colman PM: **Structure of the haemagglutinin-neuraminidase from human parainfluenza virus type III**. *J Mol Biol* 2004, **335**:1343-1357.

44. Schauer R: **Achievements and challenges of sialic acid research**. *Glycoconj J* 2000, **17**:485-499.

45. Varki A, Schauer R: **Sialic acids**. In *Essentials of Glycobiology*. Edited by Varki A, Cummings RD, Esko JD. Cold Spring Harbor Press; 2009.

46. Campanero-Rhodes MA, Smith A, Chai W, Sonnino S, Mauri L, Childs RA, Zhang Y, Ewers H, Helenius A, Imberty A *et al.*: **N-glycolyl GM1 ganglioside as a receptor for simian virus 40**. *J Virol* 2007, **81**:12846-12858.

47. Liu Y, Palma AS, Feizi T: **Carbohydrate microarrays: key developments in glycobiology**. *Biol Chem* 2009, **390**:647-656.

48. Chandrasekaran A, Srinivasan A, Raman R, Viswanathan K,
• Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R: **Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin**. *Nat Biotechnol* 2008, **26**:107-113.
In this report, the authors show that a characteristic structural topology — not the α2–6 linkage itself — enables specific binding of the influenza virus hemagglutinin to α2–6 sialylated glycans and that recognition of this topology may be critical for adaptation of hemagglutinin to bind glycans in the upper respiratory tract of humans.

49. Stehle T, Harrison SC: **Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments**. *Structure Fold Des* 1996, **4**:183-194.

50. Sauter NK, Bednarski MD, Wurzburg BA, Hanson JE, Whitesides GM, Skehel JJ, Wiley DC: **Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study**. *Biochemistry* 1989, **28**:8388-8396.

51. Arnberg N, Edlund K, Kidd AH, Wadell G: **Adenovirus type 37 uses sialic acid as a cellular receptor**. *J Virol* 2000, **74**:42-48.

52. Kitov PI, Sadowska JM, Mulvey G, Armstrong GD, Ling H, Pannu NS, Read RJ, Bundle DR: **Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands**. *Nature* 2000, **403**:669-672.

53. von Itzstein M: **The war against influenza: discovery and development of sialidase inhibitors**. *Nat Rev Drug Discov* 2007, **6**:967-974.

54. Matsubara T, Onishi A, Saito T, Shimada A, Inoue H, Taki T, Nagata K, Okahata Y, Sato T: **Sialic acid-mimic peptides as hemagglutinin inhibitors for anti-influenza therapy**. *J Med Chem* 2010, **53**:4441-4449.

55. Spjut S, Qian W, Bauer J, Storm RJ, Frangsmyr L, Stehle T, Arnberg N, Elofsson M: **A potent trivalent sialic acid inhibitor of adenovirus type 37 infection of human corneal cells**. *Angewandte Chemie* 2011, **50**:6519-6521.

56. Childs RA, Palma AS, Wharton S, Matrosovich T, Liu Y, Chai W,
•• Campanero-Rhodes MA, Zhang Y, Eickmann M, Kiso M *et al.*: **Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray**. *Nat Biotechnol* 2009, **27**:797-799.
Using glycan array screening, the authors have compared the receptor-binding characteristics of two isolates of the novel pandemic H1N1 influenza virus with those of established influenza H1N1 virus strains. The paper presents an exemplary application of glycan microarray technology.

Structure-Function Analysis of *Arabidopsis thaliana* Histidine Kinase AHK5 Bound to Its Cognate Phosphotransfer Protein AHP1

Johannes Bauer[a], Kerstin Reiss[a], Manikandan Veerabagu[b], Michael Heunemann[b], Klaus Harter[b,1] and Thilo Stehle[a,c,1]

a Interfaculty Institute of Biochemistry, University of Tübingen, D-72076 Tübingen, Germany
b Center for Plant Molecular Biology, University of Tübingen, D-72076 Tübingen, Germany
c Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

ABSTRACT   The multi-step phosphorelay (MSP) system defines a key signal transduction pathway in plants and many eukaryotes. In this system, external stimuli first lead to the activation of a histidine kinase, followed by transfer of a phosphoryl group from the receiver domain of the kinase (HK_{RD}) to downstream, cytosolic phosphotransfer proteins (HPs). In order to establish the determinants of specificity for this signaling relay system, we have solved the first crystal structure of a plant HK_{RD}, $AHK5_{RD}$, in complex with one of its cognate HPs, AHP1. AHP1 binds $AHK5_{RD}$ via a prominent hydrogen bond docking ridge and a hydrophobic patch. These features are conserved among all AHP proteins, but differ significantly from other structurally characterized prokaryotic and eukaryotic HPs. Surface plasmon resonance experiments show that $AHK5_{RD}$ binds to AHP1-3 with similar, micromolar affinity, consistent with the transient nature of this signaling complex. Our correlation of structural and functional data provide the first insight, at the atomic level as well as with quantitative affinity data, into the molecular recognition events governing the MSP in plants.

Key words: multi-step phosphorelay; phosphotransfer protein; plant signaling; sensor histidine kinase; two-component system.

INTRODUCTION

The multi-step phosphorelay (MSP) system of plants has evolved from the prokaryotic two-component system (TCS), which allows organisms to sense and respond to changes in environmental conditions. In its most basic form, the TCS consists of a membrane-bound histidine kinase (HK) that senses a specific environmental stimulus and a corresponding response regulator (RR) that mediates the cellular response, for example via the differential expression of its target genes (Laub and Goulian, 2007). Compared with most prokaryotic TCS, the MSP signaling system of plants employs a more sophisticated His-Asp-His-Asp phosphorelay system by incorporating additional signaling domains and phosphotransfer proteins (AHPs) (Figure 1A and B).

Recently, substrate specificity determining amino acid residues have been identified in the *Escherichia coli* TCS, namely for the HK-RR pairs of EnvZ–OmpR (Skerker et al., 2008; Capra et al., 2010) and $CheA_3$–$CheY_6$ (Bell et al., 2010). However, this phosphorelay in the prokaryotic TCS corresponds to an *intra*molecular phosphoryl-transfer reaction in the MSP system (Figure 1A and B). The determinants of the subsequent phosphorelay in MSP, the *inter*molecular phosphoryl transfer from the receiver domain of the kinase (HK_{RD}) to the cognate AHP, are therefore not known, and it is furthermore not understood whether and how this interaction confers specificity onto a signal.

The *Arabidopsis thaliana* genome encodes eight canonical HKs (AHK1–5, ETR1, ERS1, and CKI1), five canonical AHPs (AHP1–5), and one pseudo-AHP (AHP6) that carries an asparagine instead of the critical histidine residue (Grefen and Harter, 2004). MSP signaling regulates a wide variety of key processes in *A. thaliana*, including osmoregulation (AHK1) (Urao et al.,
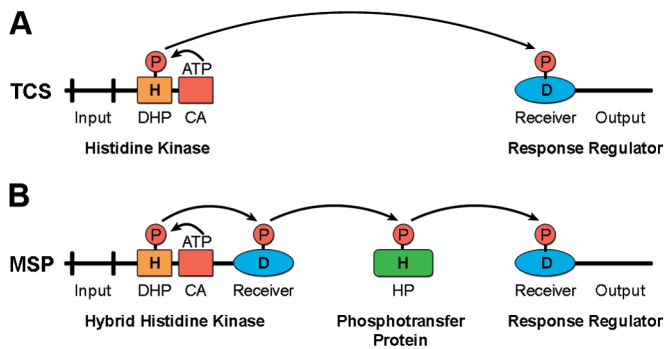
A

B

Figure 1. Schematic Overview of His–Asp Phosphorelay Systems.
(A) Overview of phosphorelay signaling in two-component systems (TCS). Signal perception induces ATP-binding to the catalytic domain (CA), followed by phosphorylation of a conserved histidine residue in the dimerization and phosphotransfer domain (DHP) of the kinase. Finally, the phosphoryl group (P) is relayed onto a conserved aspartic acid residue in the receiver domain of the cognate response regulator (RR).
(B) Overview of phosphorelay signaling in multi-step phosphorelay systems (MSP). MSP signaling involves a hybrid histidine kinase containing a fused receiver domain and an additional AHP protein.

1999; Tran et al., 2007), hormonal signaling (AHK2–4, ETR1, ERS1) (Hass et al., 2004; Nishimura et al., 2004), and mega-gametogenesis (CKI1) (Pischke et al., 2002; Deng et al., 2010; Horak et al., 2011). Recently, AHK5 was found to positively regulate salt sensitivity and to contribute to resistance to the bacterium *Pseudomonas syringae* and the fungal pathogen *Botrytis cinerea* (Pham et al., 2012). Additionally, AHK5 mediates stomatal responses to exogenous and endogenous signals such as flagellin, which alter both the intracellular level of reactive oxygen species (ROS) and redox homeostasis (Desikan et al., 2008).

AHPs can interact with several upstream HKs and downstream response regulators (ARRs) (Mahonen et al., 2006). Thus, AHPs confer an enormous flexibility onto the MSP system and, for this reason, they play a central role in signal integration (Horak et al., 2011). For example, in the well-investigated plant cytokinin signaling pathway, AHP1–3 and AHP5 act as redundant positive regulators, whereas AHP4 is thought to function in some cases as a negative regulator (Hutchison et al., 2006). Expression of the pseudo-AHP AHP6 was also shown to be induced by cytokinin (Mahonen et al., 2006). AHP6 inhibited signal transduction from a phosphorylated HK but also competed with AHP1 and blocked the cognate ARR, thus acting as a negative regulator (Mahonen et al., 2006). Notably, pseudo–AHP homologs also exist in other plant species, indicating that such negative regulation of the plant MSP system is commonly used (Mahonen et al., 2006). However, neither the structural basis of AHP-recognition nor to what extent this redundancy in AHP-recognition manifests in a similar mode of recognition have been defined.

Since the MSP system plays a central role in plant signaling, many MSP proteins have been functionally investigated in *Arabidopsis* and other plants. However, only a small number of components of the *Arabidopsis* MSP system have been structurally characterized to date: (1) the AHK4 sensor domain in complex with its cytokinin ligand (Hothorn et al., 2011), (2) the receiver domains of ETR1 (ETR1$_{RD}$) (Muller-Dieckmann et al., 1999) and CKI1 (CKI1$_{RD}$) (Pekarova et al., 2011), and (3) the DNA-binding motif of ARR10 (Hosoda et al., 2002). Importantly, structural information about the mode of transfer of the phosphoryl group from the HK to the AHP is lacking.

In order to define the specificity and mode of recognition for this interaction, we have determined the first crystal structure of a plant HK$_{RD}$, AHK5$_{RD}$, in complex with one of its cognate HPs, AHP1. This structure identifies key contacts, including a conserved AHP surface that is recognized by AHK5. In combination with *in vitro* binding studies and functional complementation assays, our analysis provides an excellent basis for understanding the molecular recognition events governing MSP in plants.

RESULTS

Crystal Structure of AHK5$_{RD}$ in Complex with AHP1

In order to reveal the structural features that underlie the AHK5–AHP interaction, we determined the crystal structure of a representative complex, between the receiver domain of the kinase, AHK5$_{RD}$, and AHP1 at 1.95-Å resolution (Table 1).

AHP1 folds into an elongated, all α-helical bundle formed by six α-helices (Figure 2A). Four of these (α3, α4, α5, and α6) form a central core structure that is augmented by helices α1 and α2. A search for structural homologs using the DALI server (Holm and Rosenstrom, 2010) revealed that the overall fold of AHP1 is most similar to those of the plant HPs MtHpt1 (pdb-code: 3US6), ZmHP2 (pdb-code: 1WN0 (Sugawara et al., 2005)), and AK104879 (pdb-code: 1YVI), with Z-scores of 23.9, 21.2, and 21.1 and r.m.s.d. values of 1.1 Å, 1.2 Å, and 1.1 Å, respectively. The conserved residue carrying the phosphoryl group, His79, protrudes from one side of the bundle, facing towards its interaction partner (Figure 2A and B).

The AHK5$_{RD}$ structure resembles other receiver domains of TCSs from different organisms (Volz and Matsumura, 1991; Volz, 1993; Muller-Dieckmann et al., 1999; Xu et al., 2003; Pekarova et al., 2011), all of which belong to the CheY-like protein superfamily (Wilson et al., 2009). Members of this superfamily possess a central five-stranded parallel β-sheet that is surrounded by five helices, giving rise to an (α/β)$_5$ topology. However, AHK5$_{RD}$ contains six instead of the usually observed five α-helices. The extra helix α4 lies on the surface opposite to the binding interface, a region in which AHK5$_{RD}$ carries about 25 additional residues compared to ETR1$_{RD}$ and CKI1$_{RD}$ (Muller-Dieckmann et al., 1999; Pekarova et al., 2011), or the yeast osmoregulator SLN1$_{RD}$ (Posas et al., 1996; Xu et al., 2003; Zhao et al., 2008). The residue carrying the phosphoryl group, Asp828, is located at the tip of strand β3 (Figure 2A and B, and Supplemental Figure 1). Loop L5, which connects strand

Table 1. Data Collection and Refinement Statistics.

AHK5$_{RD}$–AHP1	Native	NaI-soak			
	PX III (SLS)	ID 14–4 (ESRF)			
Data collection					
Wavelength (Å)	1.0000	1.2000			
Space group	P 2$_1$3	P 2$_1$3			
Cell dimensions					
a=b=c (Å)	106.8	106.6			
$\alpha=\beta=\gamma$ (°)	90.00	90.00			
Resolution (Å)	47.76–1.95 (2.00–1.95)	47.67–2.90 (2.98–2.90)			
R_{meas} (%)	6.5 / (61.0)	11.8 (67.4)			
I/σI	16.7 (2.8)	48.7 (11.5)			
Completeness (%)	99.9 (99.4)	100.0 (100.0)			
Redundancy	13.4 (11.2)	76.6 (76.7)			
Wilson factor (Å²)	40.6	52.4			
Phasing					
Sites per ASU	–	2			
Resolution bin (Å)	–	37.76–8.30	8.30–5.94	5.94–4.87	4.87–4.23
Figure of merit (%)	–	74.3	60.5	50.6	37.5
Phasing Power, isomorphous	–	1.74	1.38	(0.82)	(0.49)
Phasing Power, anomalous	–	1.87	1.64	1.21	(0.92)
R$_{cullis}$ (%), isomorphous	–	75.4	(86.8)	(88.9)	(94.5)
R$_{cullis}$ (%), anomalous	–	55.8	61.8	75.6	(82.0)
Refinement					
Resolution (Å)	1.95				
No. of unique reflections	28 327 (2063)				
R_{work}/R_{free} (%)	17.2/19.7				
No. of non-H atoms					
AHK5$_{RD}$/AHP1/water/Mg^{2+}	1052/1229/181/1				
B-factors (Å²)					
AHK5$_{RD}$/AHP1/water/Mg^{2+}	36.3/43.5/42.4/29.2				
R.m.s. deviations					
Bond lengths (Å)	0.013				
Bond angles (°)	1.248				

Phasing statistics were calculated before density modification as defined in autoSHARP (Vonrhein et al., 2007).
$R_{meas} = \Sigma_h (N/N{-}1)^{1/2} \Sigma^N_{i=1}|I_{h,i}{-}\bar{I}_h|/\Sigma_h \Sigma^N_{i=1}I_{h,i}$, where *N* is the redundancy.
$R_{work} = \Sigma_h \notin_{free} |F_{obs}{-}kF_{calc}|/\Sigma_h \notin_{free}$ Fobs
$R_{free} = \Sigma_h \in_{free} |F_{obs}{-}kF_{calc}|/\Sigma_h \in_{free}$ Fobs

β3 to helix α3, bears a γ-turn motif (residues 831–836) that is structurally conserved among receiver domains (Volz, 1993). The γ-turn is inserted into a longer loop, termed the γ-loop (Volz, 1993), which has the sequence MPVLDG in AHK5$_{RD}$. Loop L8, which connects strand β4 and helix α5 and can adopt different conformations depending on Mg^{2+}-binding and the state of phosphorylation (Guhaniyogi et al., 2006; Zhao et al., 2008), lies in close proximity to the γ-loop.

Architecture of the AHK5$_{RD}$–AHP1 Interface

AHK5$_{RD}$ and AHP1 form a 1:1 complex in solution, and this complex is also found in the crystals. Contacts between AHK5$_{RD}$

and AHP1 shield a surface of 787 Å² from solvent (Krissinel and Henrick, 2007). This surface is small in size compared with the contact areas of other protein–protein complexes (Bahadur and Zacharias, 2008). In the complex, AHK5$_{RD}$-helix α1 is inserted into a groove formed by three AHP1-helices (α2, α3, and α4) and one adjacent loop (L2) (Figure 2B). This results in a rectangular contact area, with one half primarily involving hydrophobic contacts and the other half mainly consisting of polar residues that line a ridge on AHP1 and engage in hydrogen bond formation with AHK5 (Figure 2C and Supplemental Table 1). The contact area is largely contiguous and devoid of solvent. The histidine carrying the
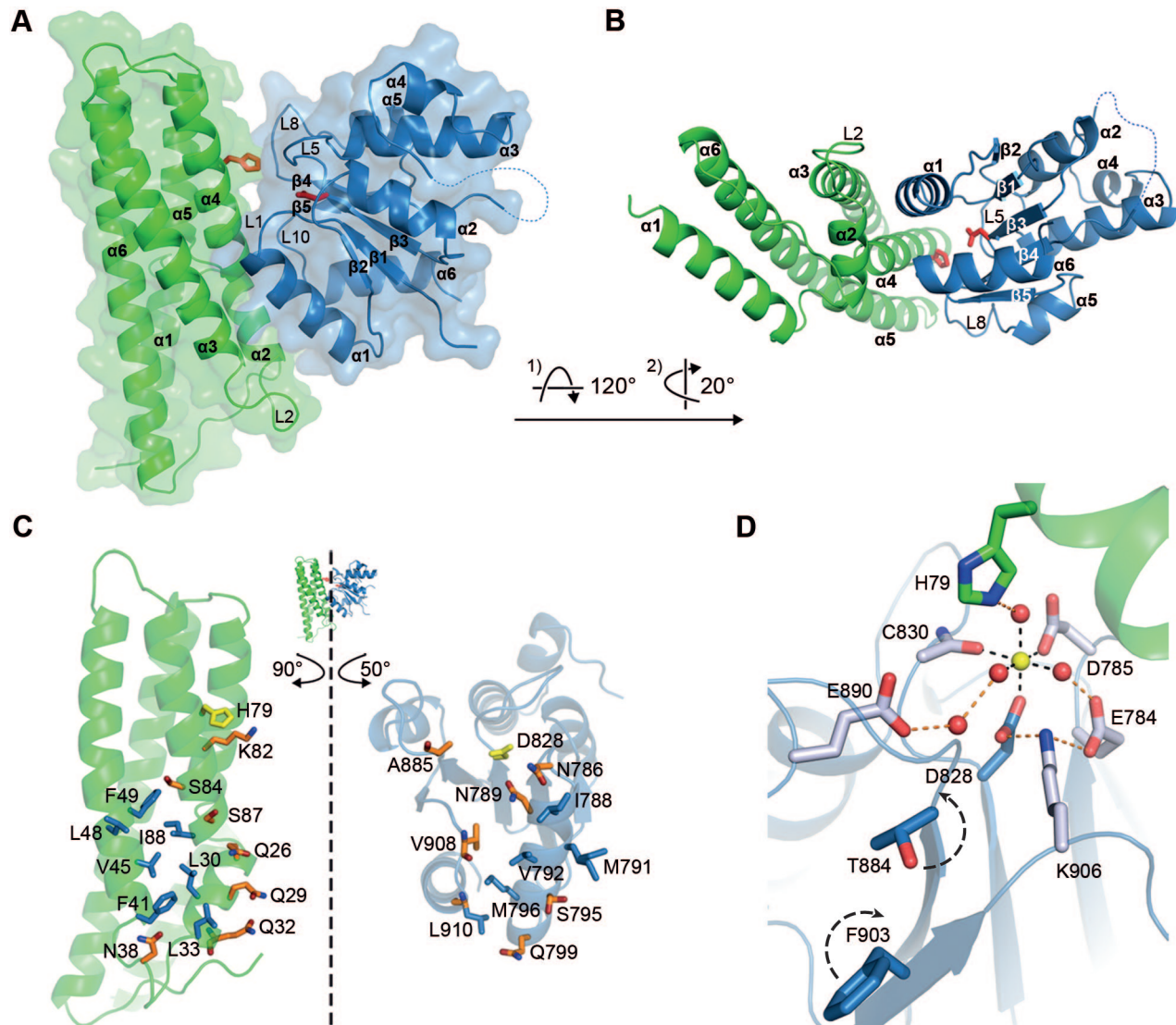
Figure 2. Structure of the AHK5$_{RD}$–AHP1 Complex.

AHP1 and AHK5$_{RD}$ are shown in all figures as ribbon tracings and colored green and blue, respectively.

(A, B) Overall structure. The conserved Asp and His residues are drawn as red sticks. In (A), AHP1 and AHK5$_{RD}$ are additionally shown in surface representation.

(C) Binding interface of AHP1 (left) and AHK5$_{RD}$ (right). Contacting residues are shown in stick representation. Residues involved in hydrogen bonds are highlighted orange, with oxygens in red and nitrogens in blue. Hydrophobic patch residues are highlighted blue. The histidine and aspartic residues that donate and accept the phosphoryl group are shown as yellow sticks.

(D) Mg^{2+}-coordination in the phosphorylation site. The Mg^{2+}-ion (yellow) and water molecules (red) are shown as spheres. The octahedral coordination geometry of Mg^{2+} is indicated by black broken lines; further hydrogen bonds of Mg^{2+}-coordinating water molecules are indicated by orange broken lines. The phosphoryl-transferring histidine and aspartic acid residues are shown as sticks, with oxygens in red and nitrogens in blue. All other Mg^{2+}-contacting residues are drawn in stick representation with carbons colored light blue. F903 and T884 of AHK5$_{RD}$ are involved in 'F-T'-coupling and are shown as blue sticks; the predicted reorientation upon phosphorylation is indicated by black arrows.

phosphoryl group, His79, lies at the very edge of the contact area, and is mostly accessible to solvent.

Phosphorylation Site of the AHK5$_{RD}$–AHP1 Complex

A Mg^{2+}-ion is located at the edge of the AHP1–AHK5$_{RD}$ interface, in close vicinity to the two residues that are involved in phosphoryl transfer (His79 and Asp828). The magnesium ion is located in an acidic pocket on AHK5$_{RD}$, and it is held in place

by three amino acids (Asp828, Cys830, and Asp785) and three water molecules that together form an almost perfect octahedral coordination sphere (Figure 2D and Supplemental Figure 2). The positions of the three water molecules are secured by either a second hydrogen bond to neighboring AHK5$_{RD}$-residues or by interactions with an adjacent water molecule. One such water molecule forms a bridge to His79 of AHP1, which would accept the phosphoryl group during the phosphotransfer process. As

Table 2. Summary of the SPR Measurements and the Calculated Affinities.

	AHP1	AHP2	AHP3
Concentration range (μM)	0.05–110.60	0.08–39.30	0.05–96.40
R_{max} (RU)	69.3	26.5	82.4
K_D (μM)	4.1	2.7	4.4
X^2 (RU)^2	0.55	0.21	0.88

the complex structure does not include a phosphoryl group, the side chain of Lys906 coordinates the carboxyl oxygens of Asp828 and Glu784. Moreover, the side chain of Thr884 points away from the phosphorylation site, so that, in turn, Phe903 is forced to occupy a solvent-exposed 'out'-conformation. We expect these two interactions to be altered in a setting in which Asp828 would be phosphorylated. As five of the six Mg^{2+}-coordination sites and all direct protein-Mg^{2+} contacts are provided by $AHK5_{RD}$, conformational changes upon Mg^{2+}-binding would primarily affect residues in $AHK5_{RD}$.

Affinity Measurements Using Surface Plasmon Resonance

In order to obtain quantitative data for the affinity and kinetics of the interaction of $AHK5_{RD}$ with AHPs, surface plasmon resonance (SPR) experiments were conducted. We immobilized anti-His$_6$ monoclonal antibodies on CM5 sensor chips, and bound purified $AHK5_{RD}$ non-covalently via its hexahistidine tag to the antibodies. Then we applied purified, untagged full-length AHP1, AHP2, and AHP3 proteins as analytes to the chip. All three proteins bound to $AHK5_{RD}$ following a 1:1 binding model (Figure 3A–F), consistent with gel filtration experiments (unpublished observations). Interestingly, the dissociation constant (K_D) values for all three proteins are highly similar *in vitro*, ranging from 2.7 to 4.4 μM (Figure 3D–F and Table 2). This indicates that the three AHPs interact with $AHK5_{RD}$ using a conserved strategy. The observed intermediate affinity is in agreement with the size of the contact area in the structure of the $AHK5_{RD}$–AHP1 complex. The association and dissociation of all $AHK5_{RD}$–AHP complexes occurred very rapidly *in vitro* and prevented us from determining association/dissociation rate constants. Rapid association and dissociation of these complexes would be required for their physiological function of quickly transferring a phosphoryl group. The remaining three AHPs (AHP4, AHP5, and AHP6) could not be purified in soluble forms and were therefore not investigated.

Specificity of $AHK5_{RD}$ for Different AHPs

To substantiate the biochemical data and to determine whether the six known phosphotransfer proteins (AHP1–6) from *A. thaliana* differ in their binding specificity for $AHK5_{RD}$ *in planta*, we conducted bimolecular fluorescence complementation (BiFC) experiments (Walter et al., 2004; Schuetze et al., 2009). Therefore, the C-terminal YFP fragment

(*YFP-C*) was fused to the N-terminus of $AHK5_{RD}$ generating *YFP-C::AHK5$_{RD}$*, whereas the N-terminal YFP fragment (*YFP–N*) was fused to the C-terminus of the different *AHPs* generating the *AHP::YFP–N* constructs. The combined expression of the BiFC constructs in transiently transformed tobacco epidermal leaf cells revealed an interaction of $AHK5_{RD}$ with AHP1, AHP2, AHP3, AHP5, and AHP6 but no interaction with AHP4 (Figure 3G and H). Although the expression levels of investigated AHP fusion proteins are similar by visual inspection, they are not identical, and thus the intensity of the fluorescence signals provides only *semi*-quantitative information about differences in their interaction with $AHK5_{RD}$. On this basis, the strongest interaction with $AHK5_{RD}$ was found for AHP2 and AHP5 followed by AHP6, AHP3, and AHP1.

Our SPR experiments show that the affinities of the investigated $AHK5_{RD}$–AHP interactions are highly similar. However, the SPR experiments were conducted *in vitro*, and thus they give only information about this particular interaction. In contrast, the BiFC assays were carried out *in planta*, and thus the $AHK5_{RD}$–AHP interaction is analyzed in the context of complete signaling cascades. Therefore, the interaction of AHPs with cognate RRs might modify the investigated $AHK5_{RD}$–AHP interaction. Furthermore, in order to maintain specificity during signaling, there are most likely (yet unknown) additional molecular mechanisms which affect and interfere with the $AHK5_{RD}$–AHP interaction *in planta*.

Conservation of AHP Binding Interface

Our SPR experiments demonstrate that recombinantly expressed, full-length AHP1, AHP2, and AHP3 engage $AHK5_{RD}$ with similar affinity. On the other hand, BiFC experiments identify differences among AHPs with regard to their *in vivo* interactions with $AHK5_{RD}$. In order to reconcile these findings with the structure of the complex, we examined the level of conservation of residues at the interface among all six AHPs proteins (Figure 4A and Supplemental Figure 3). The underlying assumptions for our analysis are (1) that the six AHPs are likely to have similar folds, which is a reasonable assumption given their high level of sequence identity (ranging from 36% to 81%), and (2) that they all form similar contacts with AHK5, which, at least in the cases of AHP1–AHP3, is likely since all three proteins have similar dissociation constants for their interactions with $AHK5_{RD}$.

Examination of the seven residues mediating hydrogen bonds showed that three of these residues, Thr26, Ser84, and Ser87, are strictly conserved in all AHPs. In addition, the hydrogen bonds mediated by Gln29 and Gln32 are functionally conserved. Although Gln29 is sometimes replaced with a glutamic acid, both residues can serve as a hydrogen acceptor for the backbone nitrogen of Leu910 of $AHK5_{RD}$. The hydrogen bond involving Gln32 is mediated by its backbone oxygen, and similar interactions can also be made by residues at the corresponding positions in the other AHPs. Thus, residues Thr26, Gln29, Gln32, Ser84, and Ser87 form a prominent docking ridge that should exist in very similar form in all six
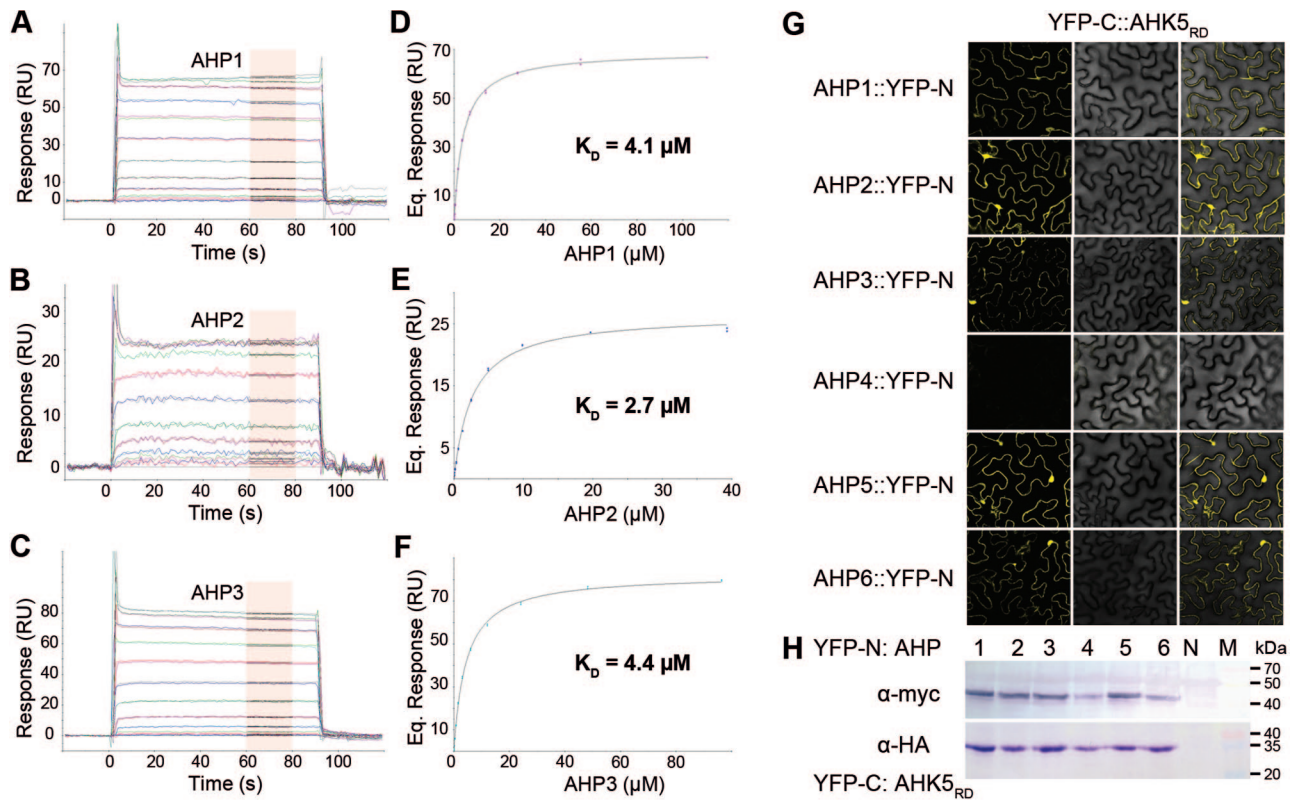
Figure 3. Interaction Studies of AHK5$_{RD}$–AHP Complexes.
(A–F) Surface plasmon resonance analyses. Double-referenced sensorgrams of the AHK5$_{RD}$-interaction with AHP1 (A), AHP2 (B), and AHP3 (C) are shown. Red boxes show the range used for calculation of averaged equilibrium response values. (D–F) Averaged equilibrium response values plotted against the injected concentrations of AHP1 (D), AHP2 (E), and AHP3 (F).
(G, H) BiFC analysis of the AHK5$_{RD}$-interaction with AHP1-6 *in planta*. (G) Confocal images of tobacco leaf cells expressing both the YFP-C::AHK5$_{RD}$ fusion protein and respective AHP::YFP–N fusion proteins. The fluorescence images (left column), corresponding bright field images (middle column), and respective overlays (right column) are shown. (H) Western blot analysis of protein extracts from tobacco leaves used for BiFC analysis. Immunodetection was carried out by α-myc antibody for AHP::YFP–N fusion proteins and by α-HA antibody for YFP::C–AHK5$_{RD}$. N, protein extract from non-transformed tobacco leaves; M, protein size standard.

AHPs, allowing them to engage AHK5$_{RD}$ using a conserved hydrogen binding pattern. Another hydrogen bond provided by the ε-amino-group of Lys82 is conserved in AHP1–5, but, as the corresponding residue in AHP6 is a valine, this hydrogen bond would not be present in an AHK5$_{RD}$–AHP6 complex. The seventh hydrogen bond is formed by the side chain amide group of Asn38 in AHP1. This asparagine is only conserved in AHP4, but is replaced with either a serine or a threonine in AHP2, AHP3, AHP5, and AHP6. Serines and threonines would not be able to form similar direct hydrogen bonds, but they would also not clash with AHK5$_{RD}$, and perhaps even allow for the formation of water-mediated, indirect hydrogen bonds.

Inspection of the hydrophobic patch revealed seven residues in AHP1 that form non-polar contacts with AHK5$_{RD}$: Leu30, Leu33, Phe41, Val45, Leu48, Phe49, and Ile88. Four of these residues (Leu30, Leu33, Phe41, and Val45) are strictly conserved in AHP1–6; the remaining three are functionally conserved with respect to their surface complementarity. Leu48 is conserved in AHP1–5, but is an isoleucine residue in AHP6. Phe49 is replaced with tyrosines in AHP4 and AHP6,

but the tyrosine hydroxyl group would be expected to point towards the AHP1-core. Ile88 is sometimes replaced with a similarly hydrophobic valine. Due to these conservative mutations, the primary characteristics of the hydrophobic patch should be maintained throughout all AHPs.

DISCUSSION

Two-component signal–response systems serve as the foundation of intracellular communication in many biological systems. Our structure-function analysis provides an understanding, at the atomic level and with quantitative affinity data, of the specificity that underlies such signaling systems in plants. The AHK5$_{RD}$–AHP1 crystal structure demonstrates that the contact area between the two proteins consists of two prominent features: the hydrogen bond docking ridge and the hydrophobic patch. Both features are highly conserved in all AHP proteins, allowing conserved modes of engagement for all known AHP proteins in AHK5$_{RD}$–AHP complexes. These findings are in good agreement with our SPR experiments,
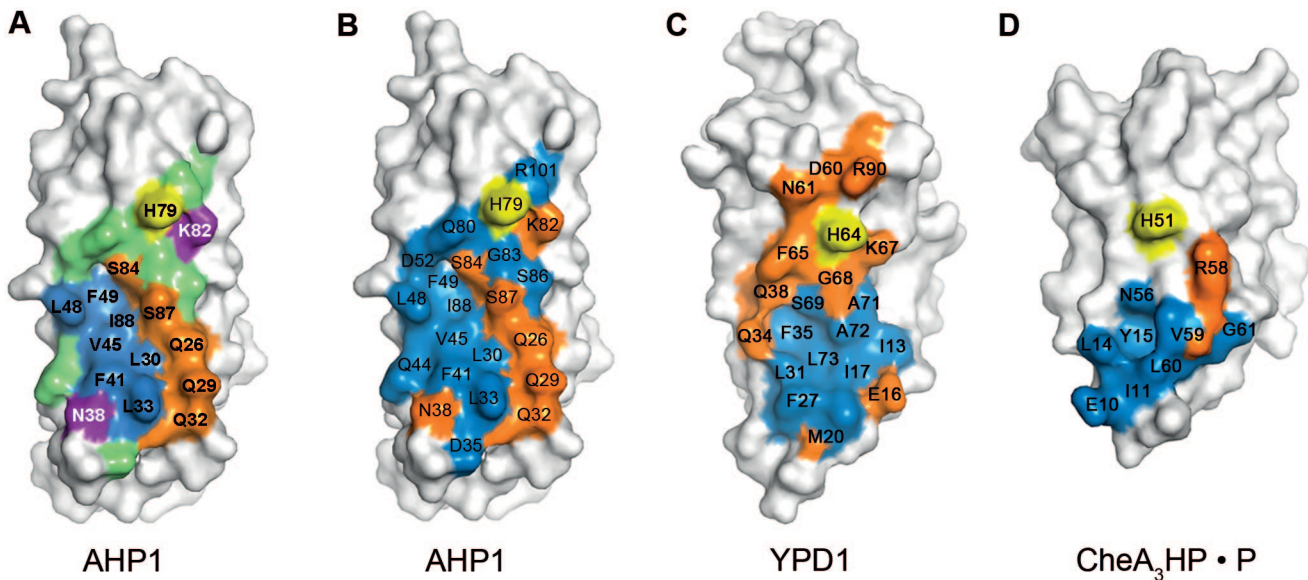
Figure 4. Interface Architectures of AHP1, YPD1, and CheA₃HP.

(A) Conservation of interface residues in AHP proteins based on the AHK5$_{RD}$–AHP1 complex structure. AHP1 is shown in surface representation with non-interface residues colored white. The phosphorylation site is shown in yellow. Functionally conserved hydrogen bonds in AHP1–6 are colored orange; not functionally conserved hydrogen bonds are colored purple. All hydrophobic residues in the interface are conserved and colored blue. All other AHP interface residues within a radius of 5 Å to AHK5$_{RD}$ are shown in green and they are all conserved in terms of their surface complementarity with respect to the interface.

(B–D) Comparison of interface architectures of AHP1 (B), YPD1 (C), and CheA₃HP (D) based on the AHK5$_{RD}$–AHP1 complex structure, the phosphorylated SLN1$_{RD}$–YPD1 complex structure, and the phosphorylated CheY₆–CheA₃HP (CheY₆–CheA₃HP•P) complex structure, respectively. All three proteins are shown in surface representation, with non-interface residues colored white. The phosphorylation site is shown in yellow. Amino acids that contribute hydrogen bonds are colored orange; residues contributing hydrophobic interactions with their respective complex partner are colored blue.

which show nearly identical binding affinities of AHK5$_{RD}$ for AHP1, AHP2, and AHP3 *in vitro*. BiFC experiments indicate that at least five of six AHP proteins can engage AHK5$_{RD}$ *in planta*, and we expect that they do so using the same contact surface and with similar affinities. The specificity of signal transduction from AHK5 to an AHP does therefore not depend on unique contacts with any one AHP, and more likely involves spatial distribution and local concentrations of an AHP. Interestingly, sequence comparison of all AHPs with other structurally characterized plant HPs from maize, rice, and clover and with numerous protein sequences found by a search in the UniProtKB database (Altschul et al., 1997) revealed that the residues involved in AHK5$_{RD}$ recognition are highly conserved across plant species (Supplemental Figure 4). Thus, the mode of AHP1 recognition serves as a template for understanding interactions in plant MSP systems.

Comparison with Non-Plant Signaling Complex Structures

Although several receiver domains ((Muller-Dieckmann et al., 1999; Pekarova et al., 2011) and pdb-codes: 3LUA, 3C97, 3GT7) and phosphotransfer proteins ((Sugawara et al., 2005) and pdb-codes: 1YVI, 3US6) have been crystallized in isolation, only two other structures of complexes between two such signal proteins have been solved to date: the osmosensing

SLN1$_{RD}$–YPD1 complex from *S. cerevisiae* (Posas et al., 1996; Xu et al., 2003; Zhao et al., 2008) and the CheY₆–CheA₃HP complex from *R. sphaeroides* (Bell et al., 2010). The receiver domains in these complexes share little sequence identity (ranging from 17% to 28%) but nevertheless possess the same fold (Figure 5A). AHK5$_{RD}$ has especially high structural similarity to SLN1$_{RD}$ (Xu et al., 2003; Zhao et al., 2008) (r.m.s.d. values of 1.0 Å), while its structural similarity to CheY₆ is less pronounced (r.m.s.d. value of 2.0 Å).

Likewise, the three HP proteins AHP1, YPD1, and CheA₃HP share the four-helix bundle core, which is assembled from helices α3, α4, α5, and α6 in AHP1 (Figure 2A and B). In contrast to both eukaryotic HPs, however, the AHP1-helix α2, which contributes three hydrogen bonds and two hydrophobic residues to the interface, is absent in prokaryotic CheA₃HP (Figure 5B). Thus, the CheY₆–CheA₃HP contact area is decreased by about 200 Å² compared to the AHK5$_{RD}$–AHP1 complex, consistent with a significant reduction in affinity of CheY₆ for CheA₃HP (K_D = 218 μM) (Bell et al., 2010). The observed differences in affinity and structure point to a central role of the AHP1-helix α2 in interface formation. As the α2 helix is also present in the SLN1$_{RD}$–YPD1 complex, this extra helix might be a common feature within the more highly evolved eukaryotic MSP systems. In addition to enlarging the contact surface, the α2 helix could also play a role in modulating interactions with selected ligands.
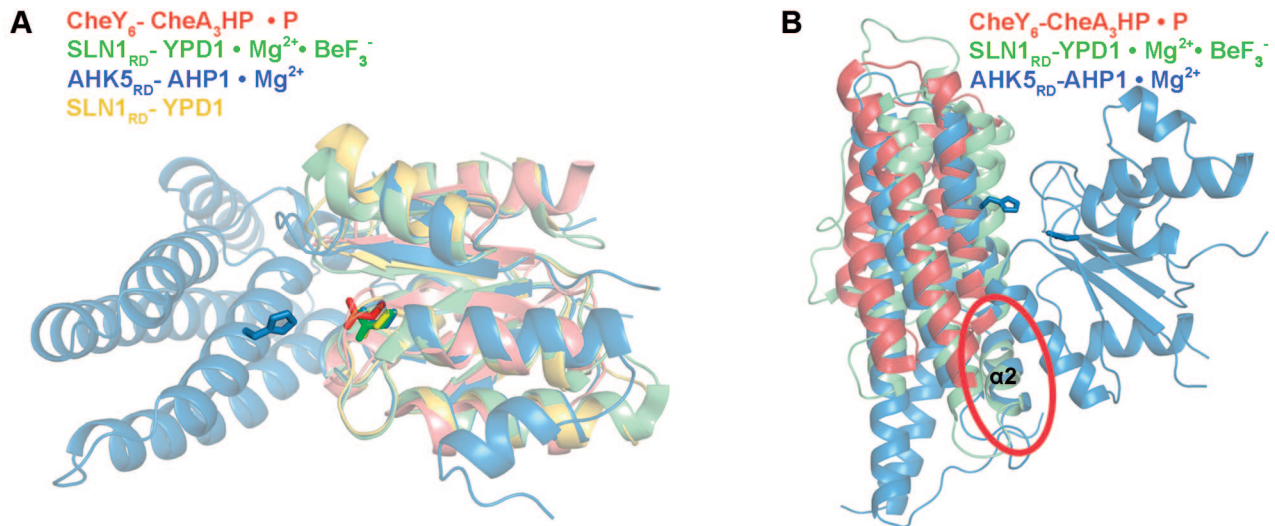
Figure 5. Superposition of AHK5$_{RD}$–AHP1 Complex with Structurally Similar Complexes.
(A) Superposition of the AHK5$_{RD}$–AHP1 complex (blue) with the phosphorylated SLN1$_{RD}$–YPD1 complex (green), the apo SLN1$_{RD}$–YPD1 complex (yellow), and the phosphorylated CheY$_6$–CheA$_3$HP complex (red) by superposing the receiver domains only. The phosphoryl-transferring histidine and aspartic acid residues are highlighted in stick representation.
(B) Comparison of the AHK5$_{RD}$–AHP1 complex (blue) with the phosphorylated SLN1$_{RD}$–YPD1 complex (green) and the phosphorylated CheY$_6$–CheA$_3$HP complex (red). The superpositions are based on the receiver domains only and, for reasons of clarity, only AHK5$_{RD}$–AHP1 is shown as a complete complex, with phosphoryl-transferring histidine and aspartic acid residues highlighted in stick representation. The extra helix of AHP1 and YPD1 is indicated.

The eukaryotic AHK5$_{RD}$–AHP1 and SLN1$_{RD}$–YPD1 complexes are well conserved with respect to the structures of their components and the overall organization of the complexes, including the sizes of the buried surface areas. It is therefore surprising that the two complexes use entirely unrelated strategies for complex formation. The amino acid residues contributing to hydrogen bonds and hydrophobic interactions are distributed differently across the interfaces (Figure 4B and C). In YPD1, nine of 11 hydrogen bond-forming residues cluster around the phosphorylatable histidine in a pattern that resembles the letter Y (Figure 4C). This Y-shaped distribution in YPD1 contrasts with the hydrogen bond docking ridge observed in AHP1 (Figure 4B). In addition, the interface in the SLN1$_{RD}$–YPD1 complex contains a hydrophobic patch that is enclosed by the hydrogen bond forming residues. Molecular recognition of CheA$_3$HP is again organized in a different manner. Only two hydrogen bonds are present in the phosphorylated CheY$_6$–CheA$_3$HP complex, and these residues bridge to a patch of residues involved in hydrophobic interactions (Figure 4D).

Structural Changes upon Phosphorylation of the AHK5$_{RD}$–AHP1 Complex

The impact of phosphorylation on the structures of HK or RR has been most extensively investigated for the bacterial chemotaxis protein CheY (Guhaniyogi et al., 2006). These studies revealed that phosphorylation of the conserved aspartate residue induces local and medium-distance changes of conserved amino acids within CheY, known as 'Y–T' coupling. In eukaryotes, this 'Y–T'-coupling was also observed in the

crystal structures of the SLN1$_{RD}$–YPD1 complex in both phosphorylation states (Xu et al., 2003; Zhao et al., 2008).

Our efforts to produce complexes of AHK5$_{RD}$–AHP1 with ions such as BeF$_3^-$ or AlF$_4^-$, two compounds that mimic the phosphorylated state of AHK5$_{RD}$, were not successful. We furthermore attempted to produce a complex with AlF$_3$, which mimics the transition state of the phosphoryltransfer reaction (Xu et al., 1997), also without success. In order to predict the structural changes that would be triggered by phosphorylation of Asp828, we therefore have to rely on comparisons with the homologous SLN1$_{RD}$–YPD1 structure (Xu et al., 2003; Zhao et al., 2008). All phosphoryl-coordinating residues are conserved between AHK5$_{RD}$ and SLN1$_{RD}$, so that the phosphoryl group would likely be coordinated by the same residues. Therefore, we predict that, upon phosphorylation, the side chain of Thr884 rotates into the active site to coordinate the first oxygen of the phosphoryl group. This would in turn make room for the Phe903 side chain, allowing it to flip from a solvent-exposed 'out'-conformation into a buried 'in'-conformation (Figure 2D).

The coordinated movement of this proposed 'F-T'-coupling in phosphorylated AHK5$_{RD}$ would result in optimal charge compensation and in a moderate rearrangement of the Thr884-containing loop L8, similar to what has been seen in the SLN1$_{RD}$–YPD1 complex structures.

Structural Implications for the Phosphoryl-Transfer Mechanism

In order to enable a phosphoryl transfer from a histidine to an aspartic acid in TCSs, two mechanisms have to be considered

(Knowles, 1980): an associative S_N2-mechanism and a dissociative S_N1-mechanism. Recently characterized TCS-complex structures favor the associative S_N2-mechanism (Varughese et al., 2006; Zhao et al., 2008). Both mechanisms require a linear O_{Asp}–P–N_{His} geometry, but they differ in their respective O_{Asp}–P and P–N_{His} distances. The dissociative mechanism involves the formation of a monomeric metaphosphate, which in turn requires that the O_{Asp}–P and P–N_{His} atom radii must not intersect, resulting in O_{Asp}–P and P–N_{His} distances of longer than 3.3 Å. In contrast, the associative mechanism requires a penta-coordinated transition state, with O_{Asp}–P and P–N_{His} distances shorter than 3.3 Å (Varughese et al., 2006).

In the non-phosphorylated $AHK5_{RD}$–AHP1 complex, the phosphoryl-donating oxygen atom of Asp828 (O_{D828}) and the phosphoryl-accepting nitrogen atom of His79 (N_{H79}) are 6.4 Å apart. The O–P–N angle α_{OPN} is 155°, indicating that O_{D828} and N_{H79} are not yet in an optimal in-line position for phosphoryl transfer. Upon phosphorylation, the Thr884-containing loop L8 in $AHK5_{RD}$ needs to be rearranged to accommodate the phosphoryl group and also to compensate for its charge. In order to provide a perfectly linear O–P–N geometry for phosphoryl transfer, His79 of AHP1 has to move about 2–3 Å closer towards $AHK5_{RD}$. One could envision that the AHK5-helix α1 serves as a hinge that immobilizes the bottom part of the complex but allows the upper part of AHP1 and especially AHP1-helix α4 to tilt towards $AHK5_{RD}$ for proper reception of the phosphoryl group. As a consequence, the O_{D828}–N_{H79} distance would decrease to about 5 Å, with optimal in-line geometry. In this scenario, the geometry in the $AHK5_{RD}$–AHP1 complex would support an associative mechanism.

Commutability of AHP Proteins

The structure-function analysis presented here extends our understanding of the parameters that underlie the specificity and affinity of plant MSP complexes. AHK5 is able to bind to a range of AHPs in a similar manner, consistent with the general commutability of AHP proteins shown for other HKs from *A. thaliana* (Mahonen et al., 2006; Deng et al., 2010). Therefore, our biochemical and structural findings are in good agreement with the hypothesis that AHP proteins can act as functionally redundant signaling hubs in the MSP system in *A. thaliana* (Schaller et al., 2008; Horak et al., 2011).

Notably, examination of conserved residues among all AHP proteins suggests that AHK5 can also bind the pseudo–AHP AHP6, but with somewhat weaker affinity. Thus, for AHK5, engagement of AHP6 might also contribute to fine-tuned signaling by blocking the MSP to certain extents depending on the cytoplasmic concentration of AHP6 relative to the concentration of canonical AHPs. Our structure and sequence analysis does not provide clues as to why AHP4 would not be able to engage $AHK5_{RD}$, nor why there are differences in the *in planta* interaction strength. We therefore think it possible that additional molecular mechanisms that interfere with the $AHK5_{RD}$–AHP complex formation, perhaps in order to fine-tune or regulate the interaction, are involved in some cases. Such mechanisms might provide additional specificity in the $AHK5_{RD}$–AHP interaction and, therefore, in MSP signaling processes in general. In combination with the established SPR protocol, the observed interactions in the $AHK5_{RD}$–AHP complex structure provide a firm platform for ongoing efforts to reengineer and evaluate MSP signaling cascades.

METHODS

Cloning

DNA coding for amino acid residues 774–922 of AHK5 ($AHK5_{RD}$) was cloned into the pET28a plasmid (Novagen) using *Nde*I and *BamH*I restriction sites. The plasmid contained an N-terminal His$_6$-tag and a thrombin protease cleavage site. Full-length AHP1–6 genes were cloned into pGEX–6PI (Amersham Biosciences) plasmids using *BamH*I and *Xho*I restriction sites. This plasmid contained an N-terminal GST-tag and a PreScission protease cleavage site.

Protein Expression and Purification

$AHK5_{RD}$ was expressed in *E. coli* Rosetta2 (DE3) cells (Novagen). Cells were grown to $OD_{595} = 0.5$ at 37°C, and the temperature was then reduced to 20°C. Gene expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Cell pellets were harvested 17 h after induction. The pellet from a 2-L culture was re-suspended in 20 mM Tris (pH 8.4), 500 mM NaCl, and 20 mM imidazole. Cells were then lysed using an EmulsiFlex homogenizer (Avestin). After centrifugation (26 000 *g*, 15 min, 4°C), the supernatant was applied onto a 1-ml HisTrap™ HP column (GE Healthcare), followed by protein elution with 250 mM imidazole. $AHK5_{RD}$-containing fractions were purified over a Superdex 75 column (GE Healthcare) using 20 mM Tris (pH 8.4), 100 mM NaCl, 5 mM MgCl$_2$, 0.5 mM TCEP, and 5% glycerol. Fractions containing monomeric $AHK5_{RD}$ were pooled, and the His-tag was then cleaved through a 40-h incubation with 15 U Thrombin protease/mg protein at 4°C. Cleaved $AHK5_{RD}$ was first applied onto a 1-ml HisTrap™ HP and then re-purified over a Superdex 75 column (GE Healthcare) using 20 mM Tris (pH 8.4), 100 mM NaCl, 5 mM MgCl$_2$, 0.5 mM TCEP and 5% (v/v) glycerol.

Expression of AHP proteins was carried out as described for $AHK5_{RD}$. In each case, the pellet from a 4-L culture was re-suspended in 50 mM Tris (pH 8.4), 150 mM NaCl, and 5% (v/v) glycerol. Cells were then lysed by EmulsiFlex homogenizer (Avestin). After centrifugation (26 000 *g*, 15 min, 4°C), the supernatant was applied onto a 5-ml GSTrap™ FF column (GE Healthcare) followed by protein elution with 10 mM reduced glutathione. The GST-tag was then cleaved through incubation for 16 h with 1 U PreScission protease/mg protein at 4°C. AHPs were further purified using a Superdex 75 column (GE Healthcare) in 20 mM Tris (pH 8.4), 100 mM NaCl, 5 mM MgCl$_2$, 0.5 mM TCEP, and 5% (v/v) glycerol. For SPR experiments, a gel filtration buffer containing 20 mM

HEPES (pH 7.5), 100 mM NaCl, and 5 mM MgCl$_2$ was used. AHP-containing fractions were applied onto a 5-ml GSTrap™ FF column for re-purification. For crystallization, AHP1 and AHK5$_{RD}$ were mixed in an equimolar ratio and then concentrated to 9.2 mg ml^{-1}.

Crystallization and Structure Determination

Crystals of AHK5$_{RD}$–AHP1 complex were grown at 20°C with the hanging drop method using a reservoir solution of 10% PEG 20 000 (w/v), 20% PEG MME 550 (v/v), 100 mM MES/imidazole (pH 6.5), and 20 mM of each following alcohol: 1,6-hexanediol, 1-butanol, (RS)-1,2-propanediol, 2-propanol, 1,4-butanediol and 1,3-propanediol. Native crystals were frozen in liquid nitrogen and used to collect data at beamline X06DA at SLS (Swiss Light Source, Villigen, Switzerland). For experimental phasing, AHK5$_{RD}$–AHP1 complex crystals were soaked for 8 min in the crystallization condition complemented with 400 mM NaI. Data were collected at beamline ID 14–4 at ESRF (European Synchrotron Radiation Facility, Grenoble, France).

Diffraction data were processed with the XDS-package (Kabsch, 2010). The structure was solved with the autoSHARP pipeline (Vonrhein et al., 2007). For initial automated model building and refinement, the ARP/wARP package (Langer et al., 2008) in CCP4 (Winn et al., 2011) was used. Structural refinement was carried out using alternating rounds of model building in COOT (Emsley et al., 2010) and restrained refinement, including TLS refinement (Painter and Merritt, 2006), with Refmac5 (Murshudov et al., 1997). Simulated annealing was performed with PHENIX (Adams et al., 2010). The final model had excellent geometry. The topology plot was prepared with TopDraw (Bond, 2003); structure figures were prepared with PyMOL (www.pymol.org, last accessed 19 December 2012). Statistics on data collection and refinement are given in Table 1. Protein sequence alignments were performed using ClustalW 2.1 (Goujon et al., 2010). Structural superpositions were calculated by Superpose using the secondary structure matching function in CCP4 (Winn et al., 2011). For O–P–N angle calculation, the unphosphorylated AHK5$_{RD}$–AHP1 complex was superimposed with the phosphorylated SLN1$_{RD}$–YPD1 complex. The position of the P-site of the phosphorylated SLN1$_{RD}$–YPD1 complex was then used for calculating the O–P–N angle in the AHK5$_{RD}$–AHP1 complex. For structural alignment, all mainchain atoms of residues 824–828 in AHK5$_{RD}$ and 1140–1144 in SLN1$_{RD}$ (corresponding to the β-strand bearing at its C-terminal end the phosphorylatable Asp residues) were superimposed using the LSQ algorithm in COOT (Emsley et al., 2010).

Surface Plasmon Resonance

SPR measurements were performed on a Biacore 2000 instrument at 13°C. Using two consecutive flow cells, an anti-His$_6$ monoclonal antibody (Novagen, N-70796) was covalently bound to a CM5 sensorchip (GE Healthcare) with the amine coupling kit (GE Healthcare). 5000–7000 RU of antibody were immobilized in both the reference and the experimental flow cells. His$_6$-tagged AHK5$_{RD}$ protein was diluted in running buffer (10 mM HEPES pH = 7.4, 150 mM NaCl, 0.05 % P20) and non-covalently bound to the monoclonal antibodies in the experimental flow cell. For AHK5$_{RD}$-binding experiments, AHP1, AHP2, or AHP3 were serially diluted in running buffer and injected for 90 s in both the reference and the experimental cell at a flow rate of 20 μl min^{-1}.

In the case of AHP1 and AHP3, 12 different concentrations ranging from 0.05 to 111 μM (AHP1) or from 0.05 to 96 μM (AHP3) were measured; 11 of these conditions were run in duplicates. For AHP2, 10 different concentrations ranging from 0.08 to 39 μM were used, and each concentration was measured twice. The data points were fitted using a '1:1 Langmuir isotherm' binding model (BIAevaluation).

To remove AHK5$_{RD}$ after each cycle, 10 μl of regeneration solution (33 mM ethanolamine, 33 mM Na$_3$PO$_4$, 33 mM piperazine, 33 mM glycine, 77 mM KSCN, 31 mM MgCl$_2$, 15 mM urea, 31 mM guanidine-HCl in the case of AHP2; 77 mM KSCN, 31 mM MgCl$_2$, 15 mM urea, 31 mM guanidine-HCl, 3 mM EDTA in the cases of AHP1 and AHP3) were injected into the measurement cell. The optimal regeneration solutions were determined through a multivariate cocktail approach (Andersson et al., 1999).

Bimolecular Fluorescence Complementation

BiFC experiments were essentially carried out as described previously (Walter et al., 2004; Schuetze et al., 2009). The cDNAs of AHP1–6 were recombined via LR-reaction into pSPYNE–35S (Horak et al., 2008), and the AHK5$_{RD}$ cDNA was inserted into pSPYCE–35S. All binary vectors were transformed into *Agrobacterium tumefaciens* strain GV3101 pMP90 and infiltrated to tobacco leaves (Horak et al., 2008). The p19 protein from tomato bushy stunt virus cloned in pBIN61 (Voinnet et al., 2000) was used to suppress gene silencing in all infiltration experiments. Abaxial epidermis of infiltrated tobacco leaves was assayed for fluorescence by confocal laser scanning microscopy (CLSM) 2 d post infiltration. Expression of the BiFC fusion proteins was determined by Western blot analysis of the transfected leaf tissues using α-HA and α-myc antibodies (Walter et al., 2004; Schuetze et al., 2009).

Accession Numbers

Atomic coordinates and structure factors have been deposited with the Protein Data Bank (accession code 4EUK).

SUPPLEMENTARY DATA

Supplementary Data are available at *Molecular Plant Online.*

FUNDING

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D–Biological Crystallography. **66**, 213–221.

Altschul, S.F., Madden, T.L., Schaeffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402.

Andersson, K., Hamalainen, M., and Malmqvist, M. (1999). Identification and optimization of regeneration conditions for affinity-based biosensor assays: a multivariate cocktail approach. Anal. Chem. **71**, 2475–2481.

Bahadur, R.P., and Zacharias, M. (2008). The interface of protein–protein complexes: analysis of contacts and prediction of interactions. Cellular Mol. Life Sci. **65**, 1059–1072.

Bell, C.H., Porter, S.L., Strawson, A., Stuart, D.I., and Armitage, J.P. (2010). Using structural information to change the phospho-transfer specificity of a two-component chemotaxis signalling complex. Plos Biol. **8**, e1000306.

Bond, C.S. (2003). TopDraw: a sketchpad for protein structure topology cartoons. Bioinformatics. **19**, 311–312.

Capra, E.J., Perchuk, B.S., Lubin, E.A., Ashenberg, O., Skerker, J.M., and Laub, M.T. (2010). Systematic dissection and trajectory-scanning mutagenesis of the molecular interface that ensures specificity of two-component signaling pathways. Plos Genetics. Plos Genetics, **6**, doi: e1001220.

Deng, Y., Dong, H., Mu, J., Ren, B., Zheng, B., Ji, Z., Yang, W.-C., Liang, Y., and Zuo, J. (2010). *Arabidopsis* histidine kinase CKI1 acts upstream of histidine phosphotransfer proteins to regulate female gametophyte development and vegetative growth. Plant Cell. **22**, 1232–1248.

Desikan, R., Horak, J., Chaban, C., Mira-Rodado, V., Witthoeft, J., Elgass, K., Grefen, C., Cheung, M.-K., Meixner, A.J., Hooley, R., et al. (2008). The histidine kinase AHK5 integrates endogenous and environmental signals in *Arabidopsis* guard cells. Plos One, **3**, e2491.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallographica Section D-Biological Crystallography. **66**, 486–501.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. **38**, W695–W699.

Grefen, C., and Harter, K. (2004). Plant two-component systems: principles, functions, complexity and cross talk. Planta. **219**, 733–742.

Guhaniyogi, J., Robinson, V.L., and Stock, A.M. (2006). Crystal structures of beryllium fluoride-free and beryllium fluoride-bound CheY in complex with the conserved C-terminal peptide of CheZ reveal dual binding modes specific to CheY conformation. J. Mol. Biol. **359**, 624–645.

Hass, C., Lohrmann, J., Albrecht, V., Sweere, U., Hummel, F., Yoo, S.D., Hwang, I., Zhu, T., Schafer, E., Kudla, J., et al. (2004). The response regulator 2 mediates ethylene signalling and hormone signal integration in *Arabidopsis*. EMBO J. **23**, 3290–3302.

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. **38**, W545–W549.

Horak, J., Grefen, C., Berendzen, K.W., Hahn, A., Stierhof, Y.D., Stadelhofer, B., Stahl, M., Koncz, C., and Harter, K. (2008). The *Arabidopsis thaliana* response regulator ARR22 is a putative AHP phospho-histidine phosphatase expressed in the chalaza of developing seeds. BMC Plant Biol. **8**, 77.

Horak, J., Janda, L., Pekarova, B., and Hejatko, J. (2011). Molecular mechanisms of signalling specificity via phosphorelay pathways in *Arabidopsis*. Current Protein & Peptide Science. **12**, 126–136.

Hosoda, K., Imamura, A., Katoh, E., Hatta, T., Tachiki, M., Yamada, H., Mizuno, T., and Yamazaki, T. (2002). Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the *Arabidopsis* response regulators. Plant Cell. **14**, 2015–2029.

Hothorn, M., Dabi, T., and Chory, J. (2011). Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4. Nature Chemical Biology. **7**, 766–768.

Hutchison, C.E., Li, J., Argueso, C., Gonzalez, M., Lee, E., Lewis, M.W., Maxwell, B.B., Perdue, T.D., Schaller, G.E., Alonso, J.M., et al. (2006). The *Arabidopsis* histidine phosphotransfer proteins are redundant positive regulators of cytokinin signaling. Plant Cell. **18**, 3073–3087.

Kabsch, W. (2010). XDS. Acta Crystallographica Section D-Biological Crystallography. **66**, 125–132.

Knowles, J.R. (1980). Enzyme-catalyzed phosphoryl transfer reactions. Ann. Rev. Biochem. **49**, 877–919.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. **372**, 774–797.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nature Protocols. **3**, 1171–1179.

Laub, M.T., and Goulian, M. (2007). Specificity in two-component signal transduction pathways. In Annual Review of Genetics (Palo Alto: Annual Reviews), pp. 121–145.

Mahonen, A.P., Bishopp, A., Higuchi, M., Nieminen, K.M., Kinoshita, K., Tormakangas, K., Ikeda, Y., Oka, A., Kakimoto, T., Helariutta, Y., et al. (2006). Cytokinin signaling and its inhibitor AHP6 regulate cell fate during vascular development. Science. **311**, 94–98.

Muller-Dieckmann, H.J., Grantz, A.A., and Kim, S.H. (1999). The structure of the signal receiver domain of the *Arabidopsis thaliana* ethylene receptor ETR1. Structure with Folding & Design. **7**, 1547–1556.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallographica Section D-Biological Crystallography. **53**, 240–255.

Nishimura, C., Ohashi, Y., Sato, S., Kato, T., Tabata, S., and Ueguchi, C. (2004). Histidine kinase homologs that act as cytokinin receptors possess overlapping functions in the regulation of shoot and root growth in *Arabidopsis*. Plant Cell. **16**, 1365–1377.

Painter, J., and Merritt, E.A. (2006). TLSMD web server for the generation of multi-group TLS models. J. Applied Crystallography. **39**, 109–111.

Pekarova, B., Klumpler, T., Triskova, O., Horak, J., Jansen, S., Dopitova, R., Borkovcova, P., Papouskova, V., Nejedla, E., Sklenar, V., et al. (2011). Structure and binding specificity of the receiver domain of sensor histidine kinase CKI1 from *Arabidopsis thaliana*. Plant J. **67**, 827–839.

Pham, J., Liu, J., Bennett, M.H., Mansfield, J.W., and Desikan, R. (2012). *Arabidopsis* histidine kinase 5 regulates salt sensitivity and resistance against bacterial and fungal infection. New Phytologist. **194**, 168–180.

Pischke, M.S., Jones, L.G., Otsuga, D., Fernandez, D.E., Drews, G.N., and Sussman, M.R. (2002). An *Arabidopsis* histidine kinase is essential for megagametogenesis. Proc. Natl Acad. Sci. U S A. **99**, 15800–15805.

Posas, F., WurglerMurphy, S.M., Maeda, T., Witten, E.A., Thai, T.C., and Saito, H. (1996). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1–YPD1–SSK1 'two-component' osmosensor. Cell. **86**, 865–875.

Schaller, G.E., Kieber, J.J., and Shiu, S.-H. (2008). Two-component signaling elements and histidyl-aspartyl phosphorelays. The Arabidopsis Book/American Society of Plant Biologists. **6**, e0112.

Schuetze, K., Harter, K., and Chaban, C. (2009). Bimolecular fluorescence complementation (BiFC) to study protein–protein interactions in living plant cells. In Methods in Molecular Biology, Pfannschmidt, T., ed. (Totowa: Human Press Inc), pp. 189–202.

Skerker, J.M., Perchuk, B.S., Siryaporn, A., Lubin, E.A., Ashenberg, O., Goulian, M., and Laub, M.T. (2008). Rewiring the specificity of two-component signal transduction systems. Cell. **133**, 1043–1054.

Sugawara, H., Kawano, Y., Hatakeyama, T., Yamaya, T., Kamiya, N., and Sakakibara, H. (2005). Crystal structure of the histidine-containing phosphotransfer protein ZmHP2 from maize. Protein Sci. **14**, 202–208.

Tran, L.-S.P., Urao, T., Qin, F., Maruyama, K., Kakimoto, T., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2007). Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*. Proc. Natl Acad. Sci. U S A. **104**, 20623–20628.

Urao, T., Yakubov, B., Satoh, R., Yamaguchi-Shinozaki, K., Seki, M., Hirayama, T., and Shinozaki, K. (1999). A transmembrane hybrid-type histidine kinase in *Arabidopsis* functions as an osmosensor. Plant Cell. **11**, 1743–1754.

Varughese, K.I., Tsigelny, I., and Zhao, H.Y. (2006). The crystal structure of beryllofluoride Spo0F in complex with the phosphotransferase Spo0B represents a phosphotransfer pretransition state. J. Bacteriol. **188**, 4970–4977.

Voinnet, O., Lederer, C., and Baulcombe, D.C. (2000). A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. Cell. **103**, 157–167.

Volz, K. (1993). Structural conservation in the CheY superfamily. Biochemistry. **32**, 11741–11753.

Volz, K., and Matsumura, P. (1991). Crystal structure of *Escherichia coli* CheY refined at 1.7-A resolution. J. Biol. Chem. **266**, 15511–15519.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol. (Clifton, NJ). **364**, 215–230.

Walter, M., Chaban, C., Schutze, K., Batistic, O., Weckermann, K., Nake, C., Blazevic, D., Grefen, C., Schumacher, K., Oecking, C., et al. (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. Plant J. **40**, 428–438.

Wilson, D., Pethica, R., Zhou, Y., Talbot, C., Vogel, C., Madera, M., Chothia, C., and Gough, J. (2009). SUPERFAMILY-sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res. **37**, D380–D386.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallographica Section D–Biological Crystallography. **67**, 235–242.

Xu, Q., Porter, S.W., and West, A.H. (2003). The yeast YPD1/SLN1 complex: insights into molecular recognition in two-component signaling systems. Structure (Cambridge). **11**, 1569–1581.

Xu, Y.W., Morera, S., Janin, J., and Cherfils, J. (1997). AlF3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP. Proc. Natl Acad. Sci. U S A. **94**, 3579–3583.

Zhao, X., Copeland, D.M., Soares, A.S., and West, A.H. (2008). Crystal structure of a complex between the phosphorelay protein YPD1 and the response regulator domain of SLN1 bound to a phosphoryl analog. J. Mol. Biol. **375**, 1141–1151.

SUMMARY

The crystal structure of the AHK5$_{RD}$–AHP1 complex reveals the determinants of specificity in plant two-component signaling. Our correlation of structural and quantitative affinity data provide the first insight into the molecular recognition events governing the multi-step phosphorelay system in plants.