The Network of Intra-HCV Protein Interactions

&

En Route to the 3D Structure of HCV E2

Dissertation

Zur Erlangung der Würde des Doktors der Naturwissenschaften des Fachbereichs Biologie, der Fakultät für Mathematik, Informatik und Naturwissenschaften, der Universität Hamburg

vorgelegt von

Nicole Hagen
aus Schwerin

Hamburg 2013

Die vorliegende Arbeit wurde in der Zeit von Dezember 2009 bis Juni 2013 unter Anleitung von Dr. Michael Schindler und Prof. Dr. Thomas Dobner und Betreuung durch PD Dr. Markus Perbandt am Heinrich-Pette-Institut – Leibniz-Institut für Experimentelle Virologie in den *Arbeitsgruppen Virus-pathogenese* und *Molekulare Virologie* angefertigt.

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

---------- ----------

Hamburg, den Unterschrift

1. Dissertationsgutachter: Prof. Dr. Thomas Dobner

2. Dissertationsgutachter: PD Dr. Markus Perbandt

Tag der Disputation: 21.06.2013

1. Disputationsgutachter: Prof. Dr. Thomas Dobner

2. Disputationsgutachter: PD Dr. Markus Perbandt

Genehmigt vom Fachbereich Biologie
der Fakultät für Mathematik, Informatik und Naturwissenschaften
an der Universität Hamburg
auf Antrag von Professor Dr. T. DOBNER
Weiterer Gutachter der Dissertation:
Priv. Doz. Dr. M. PERBRANDT
Tag der Disputation: 21. Juni 2013

Hamburg, den 04. Juni 2013

Professor Dr. C. Lohr
Vorsitzender des
Fach-Promotionsausschusses Biologie

Zusammenfassung

Nach seiner Entdeckung im Jahr 1989[87] dauerte es 16 Jahre, bis ein adäquates Hepatitis C (HCV) Zellkultursystem etabliert wurde, welches es ermöglichte den gesamten Lebenszyklus von HCV zu untersuchen[88,89]. Dies verzögerte dessen Erforschung bis zum jetzigen Zeitpunkt enorm. HCV kann den *Flaviviridae* zugeordnet werden; es ist jedoch das einzige Mitglied seines Genus Hepacivirus. Seitdem es möglich ist HCV in Zellkultur zu vermehren wurden Einblicke in die verschiedenen Abschnitte seines Lebenszyklus gewonnen. Dennoch sind die exakten Funktionen und Aufgaben der HCV Proteine bis heute noch nicht geklärt. Um das Virus und seinen Lebenszyklus besser zu verstehen, ist es wichtig das HCV Protein-Protein-Interaktions-Netzwerk aufzuklären. HCV ist in wirtschaftsstarken Staaten die Hauptursache für Lebertransplantationen. Es gibt keinen verfügbaren Impfstoff und keine ausreichend greifende Behandlung. Aus diesen Gründen ist es wichtig neue Angriffspunkte für eine anti-virale HCV-Therapie zu entwickeln.

In vorliegender Doktorarbeit benutzte ich *Fluorescence Activated Cell Sorting* (FACS; Fluoreszenz-aktivierte Zellsortierung) kombiniert mit Försters Resonanz-Energie-Transfer (FRET), um das Interaktionsnetzwerk von HCV zu entschlüsseln. Das FRET-Phänomen beruht auf dem Energietransfer von einem angeregten Spender-Fluorophor zu einem nahe liegenden Akzeptor-Fluorophor mit niedrigerer Energie. Diese Energieübertragung kann nur bei Abständen kleiner 10 nm stattfinden. Die FACS-Methode ermöglicht einen hohen Durchsatz um fluoreszierende Proben – mit wenig bis keinem Einfluss auf die Lebensfähigkeit der Zellen sowie deren Funktion – zu quantifizieren. Die Kombination beider Verfahren, FACS und FRET dient als wirksame Methode, um Protein-Interaktionen in einer Vielzahl an Zellen und Proben in einer angemessenen Zeit zu charakterisieren[90].

Alle zehn HCV Proteine wurden als Fusionen mit ECFP und EYFP konstruiert und anschließend umfangreiche FACS-FRET-Messungen durchgeführt, um potentielle Interaktionen aufzufinden. Mit Hilfe dieser Methodik wurden 20 Protein-Protein-Interaktionen mit stabilem FRET signal identifiziert. Davon konnten 12 zudem in der Leberzelllinie Huh7.5, welche den kompletten Replikationszyklus von HCV unterstützt, bestätigt werden. Sieben der gefundenen Interaktionen werden hier zum ersten Mal beschrieben. Die 13 bereits bekannten Interaktionen, konnten zudem im vorliegenden System in lebenden Zellen bestätigt werden. Des Weiteren wurden einige Interaktionen mit biochemischen Methoden untersucht, jedoch ohne Erfolg. Aus diesen Gründen – dem negativen biochemischen Nachweis und der großen Anzahl an Interaktionen, die für die verschiedenen HCV Proteine gezeigt werden konnte – spricht alles für ein dynamisches und transientes Zusammenspiel der HCV-Proteine.

Letztendlich bietet das Protein-Interaktions-Netzwerk welches hier präsentiert wird eine gute Grundlage, um komplexe Mechanismen mit Hilfe derer HCV infizierte Wirtszellen manipuliert aufzudecken.

Ein weiterer Teil der vorliegenden Arbeit war es, ein Expressionssystem zu etablieren mit Hilfe dessen die spätere dreidimensionale Strukturaufklärung der HCV-Proteine E1 und E2 (E1e AA 192-326; E2e AA 384-661) erfolgen kann. Diese Glykoproteine vermitteln das Anheften und den Eintritt des Virus in die Wirtszelle und sind die primären Angriffsziele neutralisierender Antikörper. Sie interagieren mit spezifischen Wirts-Zell-Rezeptoren und verursachen als Klasse-II-Fusions-Proteine die Fusion der Virus- und Zellmembranen. Die Zulassung eines HCV-Impfstoffes ist jedoch nicht in Aussicht. Vor diesem Hintergrund sollen zukünftig die HCV-Oberflächenproteine E1 und E2 strukturell charakterisiert werden. Vor kurzem wurde gezeigt, dass neutralisierende Antikörper, welche gegen E2 gerichtet sind, vor einer HCV-Infektion schützen können[5]. Des Weiteren können die gewonnenen Erkenntnisse dazu beitragen, bereits bestehende Daten zu erklären und neue experimentelle Ansätze zu entwickeln.

Die Reinigung von Membranproteinen stellt eine große Herausforderung dar, da hierbei spezielle Detergenzien und Methoden verwendet werden, welche in späteren Schritten die Kristallisation behindern könnten. Um dieser Problematik zu entgehen werden für das hier präsentierte Projekt die Ektodomänen der HCV-Glykoproteine benutzt. Die Expression einer sekretierten E2-Ektodomäne (E2e), welche fähig ist den Eintritt von HCV-JC1 in Huh7.5-Zellen zu behindern – ein Hinweis auf dessen Funktionalität –, wurde in einer Drosophila-Zelllinie etabliert, welche es ermöglicht große Mengen an Protein zu expremieren. Die stabil transfizierten Insektenzellen produzierten E2e zuverlässig über einen langen Zeitraum. Die E2e-Reinigung fand über dessen His-Tag mit Hilfe einer Ni-NTA-Matrix statt; weitere Maßnahmen waren Gel-Filtration und Ionenaustausch. Die Strukturaufklärung in späteren Schritten wird es ermöglichen exponierte Epitope und strukturelle Eigenschaften aufzudecken, welche hilfreich für die zukünftige Medikamenten- und Impfstoffentwicklung sind.

Summary

The Hepatitis C virus (HCV) was discovered in the year 1989[87], however, an adequate cell culture system to study HCV propagation in cell culture was not established before 2005[88,89]. This strongly hampered HCV research up to this time point. HCV was classified in the *Flaviviridae* family, but it is the only member of its genus hepacivirus. Since it has been possible to propagate HCV in cell culture, light was shed on the diverse steps of its life cycle. Nevertheless, the exact role and features of the different HCV proteins are not yet clear. Hence, elucidation of the HCV protein-protein interaction network would help to understand both the virus and its life cycle in more detail. As the main cause of liver transplantations in developed countries, with no vaccine available and only limited therapeutic treatment options, it is important to establish novel targets for anti-HCV therapy.

In this PhD thesis, I used Fluorescence Activated Cell Sorting (FACS) combined with Foersters Resonance Energy Transfer (FRET) to elucidate the interaction network of HCV proteins. The FRET phenomenon is an energy transfer from an excited donor fluorophore to a lower energy acceptor fluorophore, which only occurs at distances below 10 nm. FACS enables a high throughput measurement to quantify fluorescent samples with little or no effect on cell viability and function. Combination of both, FACS and FRET serves as a powerful tool to characterize protein interactions among large numbers of cells and samples in a feasible amount of time[90].

All ten HCV proteins were generated as fusions with ECFP and EYFP and extensive FACS-FRET experiments were conducted to elucidate their potential interactions. Using this approach 20 HCV protein-protein interactions showing a robust FRET signal were identified. 12 out of these could be confirmed in the liver cell line Huh7.5, which supports the full replication cycle of HCV. Seven of these interactions have not been described in the literature before. Conversely, 13 of the interactions reported herein were described previously and could now be additionally confirmed in living cells. Furthermore, some interactions were tested in biochemical approaches, however, without success. Thus, the absence of biochemical co-immunoprecipitation and the large number of interactions described for the various HCV proteins argues for dynamic and transient protein interplay of HCV proteins. In sum, the protein interaction network reported here represents a firm basis to elucidate the complex mechanisms by which HCV manipulates infected cells.

A second part of the project was to establish an expression system suitable to assess the 3D structure of HCV E1 and E2 (E1e AA 192-326; E2e AA 384-661). These glycoproteins mediate cellular

attachment and entry of the virus and are the primary targets of neutralizing antibodies. E1 and E2 interact with specific host cell receptors and mediate the fusion of virus and cell membrane acting as class-II fusion proteins. The approval of an HCV vaccine is urgently needed. Therefore, it is a main goal to characterize the structure of the HCV envelope proteins E1 and E2. It would be a milestone in HCV research to solve the three-dimensional structure of these glycoproteins. It was already shown that neutralizing antibodies recognizing E2, protect against an HCV infection[5]. Additionally, gained results will help to explain already existing data and to develop new experimental approaches.

Purification of membrane proteins is challenging since specific detergents and methods are needed which in turn can hamper crystallization. The presented project uses ectodomains for purification, circumventing this problem. The expressed and secreted E2 ectodomain (E2e) was able to compete with HCV-JC1 for cell entry in Huh7.5 cells, indicating its functionality. E2e was expressed using a Drosophila cell line yielding high amounts of protein. The stably transfected insect cells produced E2e over a long period without loss of productivity. E2e was purified via its His-tag using a Ni-NTA matrix, gel filtration and ion exchange. Solving the structure in later steps will enable the discovery of exposed epitopes and structural features, which could then be used for rational drug design and vaccine development.

Content

1. Introduction

Proteins operate with others in complexes and networks. Protein-protein interactions (PPIs) occur when two or more proteins bind each other to carry out their biological function. There are multi-subunit proteins, which are composed of more than one protein. Common examples are hemoglobin or the more complex bacterial flagella apparatus. PPIs regulate many important molecular processes in cells such as DNA-replication, transcription, translation, splicing, secretion, cell cycle control, signal transduction, and intermediary metabolism. Furthermore, PPIs control post-translational modifications, cell differentiation, protein folding, and transport[91]. Altered protein interactions, however, can cause different diseases. Thus, it is important to identify protein interactions and their degree of regulation, to figure out consequences of their interplay, to elucidate functions of proteins or the emergence of diseases and hence to develop new therapeutic approaches.

Protein interactions in general can be characterized as stable or transient. Both can be strong or weak, fast or slow. They can be mediated by small regions (domains), by large surfaces (e.g. leucine zipper), or by intermediate forms of these two extremes[91]. All interactions of a biological system are called 'interactome'.

Stable interactions are observed in multi-subunit complexes, where the single units can be identical or different. Subunit complexes are more common in nature since their translation is less error-prone compared to large proteins. Only subunits – and not the whole protein – have to be eliminated, if translation is somehow defective. Long lasting interactions are needed to build these complexes. However, interactions are dynamic processes, for instance the formation and re-formation of actin-filaments[92]. Stable interactions are best studied by co-immunoprecipitation, pull-down assays or Far-Western-blot analysis.

Transient interplays, e.g. during protein modification through protein kinases, phosphatases and proteases, are short interactions and in turn result in further changes of PPIs. PPIs in general are expected to control the majority of cellular processes and typically require a specific set of conditions regulating the interaction. They can be demonstrated via crosslinking or label transfer methods and are the most challenging to identify by using physical methods.

PPIs can be visualized within a network or a map. Often interaction maps are very complex and difficult to generate manually. For this purpose, online tools and software programs e.g. *Cytoscape* or *OS prey* are available to manage large datasets. In addition, databases sum up described and

predicted interactions. Protein networks provide a good starting point to study the function of individual proteins and regulatory pathways. However, additional approaches are needed to verify interactions and to test their actual function *in vivo*. A comprehensive overview of different methods to detect protein interactions is given by Berggård et al.[93]

Since proteins mainly act via interactions, PPI networks provide many insights into protein functions and dynamics. The network biology simplifies complex systems, which paves the way to a greater knowledge of biological processes. Large-scale networks are available for pairwise protein interactions as it was shown *inter alia* for human proteins by Stelzl et al.[94] and for *Saccharomyces cerevisiae* by Tarassov et al.[95] Networks can also be generated for protein complexes and genetic compounds with direct or indirect interactions, ascertained via experimental or predicting methods. Many mapping examples are listed in the review of Trey Ideker and Nevan Krogan[96]. Based on the large amount of data, several databases, for instance BioGRID and PRIME (more listed in the appendix), consolidate already demonstrated and additionally predicted interactions.

In general, viral protein interactions can be characterized as interactions of viral proteins with each other (*intra*-viral) or as the interactome of viral proteins with cellular factors of the host (*inter*-viral). Within this thesis, *intra*-viral interactions of Hepatitis C proteins were investigated and an expression system for the structural analysis of HCV E2 glycoproteins was established.

HCV emerged as a significant global health problem over the past decades. The relatively late discovery of a non-A, non-B viral hepatitis genome in the year 1989[87] – later called Hepatitis C – strongly delayed HCV research up to this time point.

Besides HIV and HBV as one of the most important viral blood-borne infections, the medical impact of Hepatitis C is severe. HCV is a serious and growing health problem throughout the world: Annually about four million people get infected by HCV and tens of thousands die[97]. Furthermore, hepatocellular carcinoma (HCC) is one of the most common and life-threatening tumors in the world[98] and HCV-related chronic liver diseases will affect four times more people in 2015 than it did in 1990[99,100]. Hepatitis C is a significant social burden, since billions of dollars were invested in healthcare costs[100], and since it is the main reason of liver transplantations in developed countries[101]. It causes severe chronic hepatitis (in 75 – 90 % of the patients[102]), which often leads to liver cirrhosis (in 10 – 40 % of infected patients[103]). The most severe complication of HCV infection is the development of hepatocellular carcinoma (HCC). About 80 % of HCV infections are characterized by a progression of the disease. Transmission occurs via injection drug users, transfusion of infected blood

products, and inappropriate use of needles and syringes. Sexual transmission has been suggested as well[104].

Worldwide 130 to 210 million people are infected with HCV, this corresponds to 3 % of the world population; 350,000 of which die per year. The numbers of infected patients decreases in the US, Europe and Japan, but in the developing countries HCV infection still is on the rise[105,106]. Most recent studies indicate that HCV might be more deadly than HIV[107]; in comparison, the annual funding of HIV research in the US is 30 times higher than that for HCV, as depicted by a report on the NIH homepage for the years 2007 to 2012 (http://report.nih.gov/categorical_spending.aspx).

Since HCV and HIV share transmission routes, co-infections of HIV and HCV are very common. Three percent of the HCV-infected persons are additionally infected with HIV, which has a negative prognosis on the course of HCV infection. About 20 to 30 % of HIV patients are co-infected with HCV in the US. Whether HCV has a negative effect on the course of HIV infection is not clear yet[108]. The main morbidity in co-infected patients is caused by liver related deaths, primarily due to HCV. Co-infection rates among injection drug users with prevalence rates up to 90 % are very high[109].

At present, there is no preventive vaccine against HCV available. The last decade, treatment mainly consists of a combination therapy of pegylated interferon-α and Ribavirin. Ribavirin interferes with RNA metabolism required for viral replication. The specific mechanisms are still unknown. Interferon, on the other hand, degrades cellular and viral RNA and boosts adaptive and innate immunity, e.g. generation of MHC class-I molecules and activation of natural killer cells. Early treatment of acute HCV infection with interferon-α (IFN- α) is the only therapeutic option to prevent a chronic course of HCV infection. However, IFN-α treatment is very expensive, has many adverse effects and works only in half of the treated patients[102]. The worst side effects are anemia, depression, pain in joints and muscles, insomnia and loss of neutrophiles, which leads to a reduced response to bacterial infections.

New promising therapy approaches are Boceprevir (Merck) and Telaprevir (Vertex Pharmaceuticals)[110,111]. These drugs are HCV NS3/4A protease inhibitors, which have successfully passed clinical phase III trials. They were used in combination with interferon-α and Ribavirin and were effective in treating genotype 1 infected patients as well, which resisted standard therapy. Thus, combination treatment could be the new standard therapy approach, since genotype 1 is the most spread in the world. These new drugs cannot replace Ribavirin + Interferon-α treatment and combination therapy results in more severe adverse effects in patients. Therefore, additional

therapeutic regimens are needed, which should involve interferon free therapies. Some phase II trials are currently testing the combination of NS3/4A, NS5A or NS5B inhibitors[112].

1.1.1 HCV replication cycle

Members of the *Flaviviridae* possess a positive-sense, single-stranded, enveloped RNA genome. Compared to the flavivirus genus, the remaining groups of *Flaviviridae* (hepacivirus, pestivirus and the group of GB virus C) have an internal ribosomal entry site (IRES) instead of a type-I cap structure at the 5'-UTR.

HCV is the only member of its genus hepacivirus with six identified genotypes. These genotypes can be divided in many subtypes. Genotype 1 is the most common; genotype 2 is not as common, but mainly distributed in industrialized countries. Genotype 3 mainly appears in intravenous drug users and genotypes 5 and 6 are not very frequently found.

HCV enters hepatocytes via the interaction of its two envelope proteins E1 and E2 with four known host receptors (CD81, SRB-I, claudin-1, and occludin). After entry, host ribosomes bind to the IRES and translate the HCV polyprotein. After polyprotein processing, oligomerization of NS4B distorts the host ER into a membranous web comprising the HCV replication complexes. Alteration of cellular membranes is characteristic for positive-strand RNA viruses[113]. Vaccinia viruses, for example, use Vimentin, an element of the cell cytoskeleton. *Asfarviridae*, which cause severe disease in pigs, exploit ribosomes and polysomes as well as cytoskeletal elements and membranous material. Enveloped RNA viruses, e.g. bunyaviruses, recruit mitochondria and establish their replication complexes in tubular structures from Golgi membranes. These membranous structures, also called viroplasms, act as virus factories[114]. Within these factories, HCV RNA polymerase NS5B transcribes HCV RNA in the membranous web. Next, genomes are translocated to lipid droplets, where virion assembly takes place. Lipid droplets are highly dynamic lipid storage organelles, specific for liver and intestine cells. Similar mechanisms were also demonstrated for Dengue virus[115] and Rotavirus[116]. HCV replication takes place in hepatocytes, but in patients with HIV, viral replication can also be observed in peripheral blood mononuclear cells (PBMCs)[108,117]. Furthermore, Hepatitis RNA was detected in brain tissue[118]. Finally, HCV particles associate with VLDL (very low density lipoproteins) and lipid droplets and are most likely secreted from cells via exocytosis[119,120] to infect new cells.

1.1.2 HCV Genome Organization

The HCV polyprotein consists of about 3000 amino acids (AA). Signal sites at the C-terminus lead to polyprotein cleavage into structural (Core, E1, E2) and non-structural (p7, NS2, NS3, NS4A & B and NS5A & B) proteins. For this purpose the hosts ER signal peptidase cleaves after Core, E1, E2 and p7[113], the virus-encoded protease NS2-3 cleaves between NS2 and NS3. All proteins downstream of NS3 are cleaved by the NS3/4A serine protease (Fig. 1).

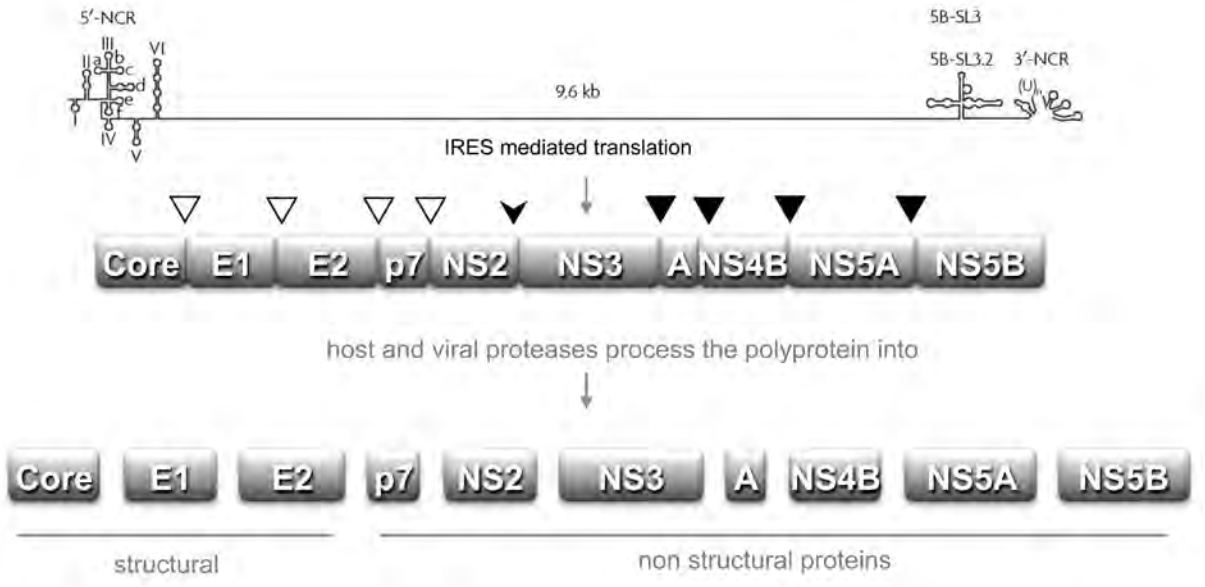

Fig. 1: **HCV genome organization and polyprotein processing by cellular and viral proteases**. The HCV RNA is depicted with its secondary structures in the 5' and 3' non-coding regions. IRES mediated translation leads to a polyprotein, which is further processed by cellular and viral proteases into structural and non-structural proteins. Core, E1, E2, and p7 are cleaved via the host ER signal peptidase (▽), while NS2 cleaves between itself and NS3 (▼). The complex NS3/4A cleaves all proteins downstream of NS3 (▼). Adapted from Moradpour et al., 2007 and Dubuisson et al., 2002.

An internal signal sequence in the Core C-terminus directs the maturing polyprotein to the host cell ER membrane and translocates the E1 ectodomain into the ER lumen. Signal peptide cleavage via the signal peptidase at the C-terminus and further processing via the signal peptide peptidase leads to mature Core protein. Afterwards the signal sequence is reorientated towards the cytosol (single transmembrane passage)[113]. The manifold functions of HCV proteins are indicated in Fig. 2.

	HCV Protein	Function			TMD	References	
structural proteins	Core	forms viral nucleocapsid; RNA binding; homo-oligomerization; interacts with lipid droplets; minor proportion present in the nucleus; attached to the ER and on the surface of lipid droplets			acts as E1 signal peptide before further processing through signal peptidase, further processing via signal peptide peptidase	1, 2, 3	
	E1	envelope glycoproteins; forming of a non-covalent complex to build the viral envelope; responsible for receptor binding and virus entry	modified by highly conserved N-linked glycans	up to six glycosylation sites	posses ER retention signals for anchoring; thought to form a hairpin structure; C-terminus can function as signal sequence, signal sequence cleavage, membrane translocation after cleavage	4	11, 12, 13, 14, 15
	E2			highly glycosylated via high-mannose glycans; up to 11 potential glycosylation sites		5, 6, 7, 8, 9, 10	
non-structural proteins	p7	essential for productive infection in vivo; forms oligomers: ion channel like structure; cation channel activity: viroporin			has two TM segments, connected via a short cytoplasmatic loop, C-terminus can function as signal sequence	16, 17, 18, 19, 20	
	NS2	essential for complete replication cycle *in vitro* and *in vivo*, C-terminal part of NS2 and N-terminal part of NS3: protease activity	NS2 essential for production of infectious virus		TMD in N-terminus	21, 22	
	NS3		N-terminal part of NS3: Serin-protease; cleaves and inactivates proteins of the innate immune system: Trif (TICAM-1) and Cardif (MAVS, IPS-1, VISA)	helicase activity in the C-terminus	anchorage in ER via NS4A	23, 24, 25, 26, 27	28, 29, 30
	NS4A	co-factor for NS3; membrane anchor for NS3			TMD in N-terminus; ER localization of NS3/4A complex		
	NS4B	induces formation of membraneous web	associated with lipid rafts, therefore responsible for anchorage of remaining HCV NS proteins in lipid rafts		contains at least four TMDs in its middle part	31, 32	
	NS5A	serine-phosphoprotein, basally phosphorylated and hyperphosphorylated forms, mediated by CKI (protein kinase); possibly accomodates RNA; involved in RNA replication			TMD in N-terminal amphipatic alpha-helix; ER localization signal	33, 34, 35, 36, 37	
	NS5B	RNA dependent RNA polymerase	phosphorylation through PKR2 can influence HCV replication		TMD in C-terminus; ER localization signal	38, 39, 40, 41, 42	

Fig. 2: **HCV proteins and its functions.** Viral proteins are often multi-functional. Therefore, they exhibit different activities. Additionally, localization of each transmembrane domain (TMD) and mode of HCV protein anchorage into the host ER membrane is indicated.

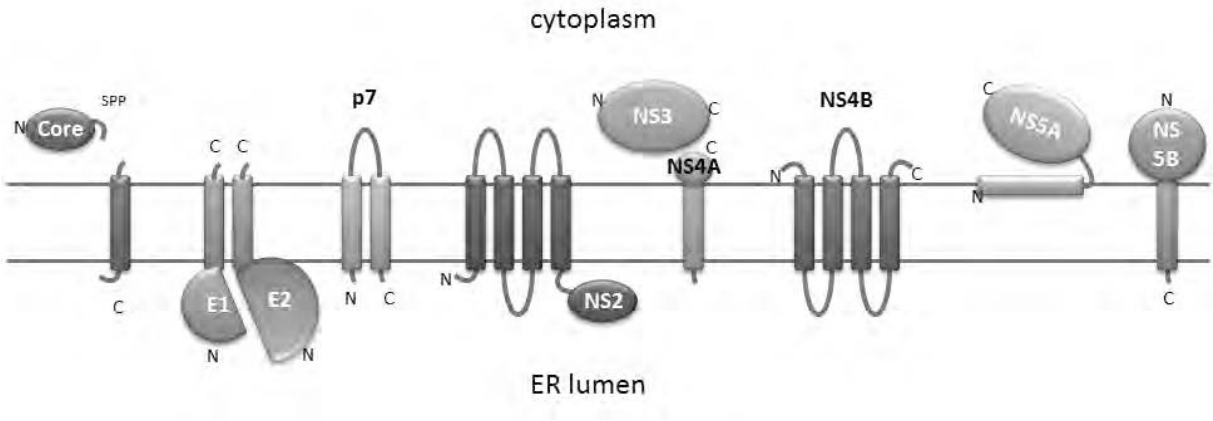

Fig. 3: **HCV proteins anchored in the host ER membrane.** Depicted is the arrangement of HCV proteins in the host ER membrane after polyprotein processing via cellular and viral proteases. N- and C-termini are indicated as N and C, respectively. Signal peptide peptidase (SPP) cleaves Core as signal peptide of E1. Moreover the interplay of the E1/E2 glycoprotein heterodimer and association of NS3 with NS4A is indicated.

1.1.3 Hepatitis C Cell Culture System

Hepatitis C exhibits a narrow host range and thus, it naturally infects humans exclusively. These aspects caused difficulties to establish an animal model. Chimpanzees can be infected with human HCV sera, although infectivity is low[97]. Due to the problematic ethical background, the high experimental costs, the low sample size and a low infectivity of human HCV sera in chimpanzees, *in vivo* alternatives are urgently needed. Therefore, Korzaya et al.[121] studied spontaneous infection in six different simian species and screened for antibodies against HCV proteins revealing that virtually all tested species contained HCV antibodies. Thus, other monkeys than chimpanzees, i.e. Old World lower primates, might be suitable as an alternative monkey model.

Some already available non-monkey systems are rodent models bearing human hepatocytes, e.g. immunocompetent fetal rats, immunodeficient trimera mice and others[99,122].

No cell culture system to study HCV propagation was available until 1999 when Lohmann et al.[123] established a replicon system based on the human hepatoma cell line Huh7. Such replicon systems are also known for LaCrosse[124], Influenza[125], and Lassa Virus[126]. Furthermore, one single cell clone of the Huh7 cell line, named Huh7.5, turned out to be more permissive for efficient HCV RNA replication than others[127]. This cell clone had a single point mutation in the intracellular receptor RIG-I, proposed to render the cells more permissive for HCV[127]. However, this hypothesis was disproved by the work of Feigelstock et al. in 2010[128]. Therefore, it is still not known why this cell clone is much more permissive than others.

Huh7.5 cells electroporated with HCV JFH1 RNA lead to the formation of infectious virus in cell culture[89]. The electroporated genome consists of full-length JFH1 genome (genotype 2a), which was

isolated from a Japanese patient with fulminant Hepatitis. A T7 promotor is inserted immediately upstream of the HCV genome for the transcription via T7 RNA polymerase. The JC1 strain is a chimera of J6 (Core to NS2) and JFH1 (NS3 to NS5B)[129] and more infectious than JFH1.

1.1.4 Immune Response Towards Hepatitis C Virus

Acute HCV infection often comes along with mild symptoms as fatigue, pain in muscles and joints, weight loss or even no symptoms, and is therefore often not recognized[130]. In acutely infected patients viral clearance is associated with a strong CD4+ and CD8+ T-cell response[99]. Chronically infected patients show high levels of neutralizing antibodies, however, HCV acquires neutralization escape mutations[131], which allow the virus to evade host immune response resulting in a persistent infection. Quasispecies can develop due to viral genome amplification by the HCV RNA polymerase, which has no proofreading. In the end, this results in related but genetically distinct viral variants, leading to high variability of HCV within a patient[132]. Interestingly, most of the genetic variations mainly take place in the hypervariable region 1 (HPV-1) located in E2 (see Fig. 8 in methods part).

HCV triggers, controls and evades hepatic host response. It stimulates the production of IFN and activates other cellular genes that could theoretically control the infection. Therefore, HCV seems somehow resistant to various antiviral pathways.

For instance, HCV NS3/4A complex blocks phosphorylation and therefore the activity of IRF3 (interferon regulatory factor 3) by inactivating RIG-I (retinoic acid-inducible gene I). RIG-I and TLR3 (toll-like receptor 3) recognize dsRNA of HCV. Double-stranded RNA arises due to regions of extensive secondary structure encoded by the HCV RNA (Fig. 1). Another example is NS5A, which activates STAT3 (signal transducer and activator of transcription 3); in turn STAT3 activates the Jak (Janus kinase)-STAT signaling pathway. Furthermore, HCV Core protein leads to an increased expression of SOCS3 (suppressor of cytokine signaling 3), which is a negative regulator and inhibitor of the Jak-STAT signaling. Many other strategies of HCV interference with host cell signaling pathways were identified[133]. In conclusion, HCV exhibits two immune evasion strategies: subversion of the interferon response and mutational escape.

1.1.5 HCV Protein Interactions

Viruses severely influence the host cell protein-network, despite their relatively small genomes. For HCV this was shown in different network screens. Noteworthy, for assessing stable and transient intra-viral PPIs, the different experimental approaches have its pros and cons, especially regarding

their sensitivity and specificity. The most common ones are Yeast two hybrid (Y2H) and Co-immunoprecipitation (Co-IP). Y2H is a very simple method, but has high rates of false positives and uses a different cellular context, which cannot provide needed protein modifications. For Co-IPs, expression takes place in the natural cell context, but specific antibodies are needed. Additionally, cell lysis is a prerequisite, which destroys previously separated compartments.

1.1.5.1 FACS-Based FRET

The above-mentioned disadvantages of Co-IPs and Y2H screens demonstrate the need of alternatives to perform protein interaction screens in living mammalian cells. In this thesis, Fluorescence Activated Cell Sorting (FACS) combined with Foersters Resonance Energy Transfer (FRET) was used, to elucidate the interaction network of HCV proteins[90]. HCV is suitable for this medium-throughput approach due to its very small genome.

In general, PPIs are important for biological functions. Given that proteins can be part of large complexes, they are often multifunctional and activated through other proteins. Protein-protein interaction maps can reveal the overall physical and functional landscape of biological systems[96]. Many issues of HCV replication or features of its proteins remain to be uncovered. Hence, elucidation of the HCV protein-protein interaction network can facilitate understanding of the virus and its life cycle in more detail. The identified interactome will encourage the detection of new target proteins for therapies and hence, the development of new drugs.

Initially, FACS-based FRET has been used to measure FRET between differentially labeled monoclonal antibodies, and was already extensively used in the 80s by the group of Szöllőssi, which introduced the term flow cytometric energy transfer (FCET). The Szöllőssi group labeled target proteins with specific fluorescent dyes[134-135] and subsequently worked with CFP and YFP chimeras[136] resembling the approach in the present work.

1.1.5.2 Flow Cytometry

Flow cytometry allows analyzing single cells in suspension due to specific light scattering and fluorescent characteristics. Fluorescence activated cell sorting, as a specialized type of flow cytometry is a high throughput measurement to quantify fluorescent samples via individual detection of each cell in a fluid stream through measuring their physical and chemical characteristics[137], with little or no effect on cell viability and function. Flow cytometers measure relative fluorescence, size, and granularity of a single cell via the interaction with a laser beam at high velocity[138]. First practical flow cytometry was performed in the 1940s to count cells and bacteria, and first fluorescence

cytometers were built in the late 1960s[137]. Although, repeated examination of one cell or studies of the same cell over time are not possible[137], the big advantage is to characterize large numbers of cells and samples in an acceptable amount of time. Furthermore, co-expressed, differently labeled proteins and to some extent, the levels of expression can be detected[137].

1.1.5.3 Foersters Resonance Energy Transfer – FRET

The FRET phenomenon – first described by Foerster in the year 1948[139], and rediscovered in the late 60s[140] – is an energy transfer from an excited donor fluorophore to a acceptor fluorophore within a radius below 10 nm. Partial overlap of the donor emission spectra with the acceptor excitation spectra is a prerequisite for FRET (Fig. 4). However, if fluorescence spectra are highly similar, distinction of FRET from the emission of the acceptor is no longer possible.

Co-localization studies with immunostained fixed samples are also used to determine PPIs. The resolution limit of about 200 nm on a microscope under ideal conditions does not give evidence for a real interaction, since the size of a typical protein is about 10 to 100 nm. Additionally, due to cell fixation, detection of dynamic processes is not possible. FCET circumvents these problems and is suitable to measure also weak or transient interactions, which are often not measurable with common techniques as Co-IP, Y2H, and pull-down approaches. For the FACS-based FRET method the construction of fusion proteins is necessary, since labeling inside living cells is not possible. Therefore, genetically encoded fluorescent proteins (CFP & YFP) fused either C- or N-terminal to the HCV proteins were used in the present work.

FRET only occurs at distances of 8 to 10 nm or even less, and detects a variety of intra-molecular interactions like protease cleavage, calcium signaling and phosphorylation[141], and inter-molecular interactions as described in this work. In addition, oligomerization of proteins and conformational changes in the same molecule can be monitored[142].

Energy transfer occurs via long-range dipole-dipole interactions. The excited fluorophore correlates to an oscillating dipole that undergoes an energy exchange with a secondary dipole with similar resonance frequency. FRET in principle does not require a fluorescent acceptor molecule, but in most cases both, donor and acceptor are fluorescent. The non-radiative energy transfer results in quenching the donor fluorescence, the reduction of its fluorescence lifetime, and simultaneously leads to an increase of acceptor fluorescence emission. FRET efficiency E_{FRET} describes the quantum yield of energy transfer transition according to the distance between two fluorophores:

$$E_{FRET} = 1/[1 + (r/R_0)^6]$$

r = distance between the two molecules

R_0 = Foerster radius: characteristic distance where FRET efficiency is 50 %, can be calculated for any pair of fluorescent molecules; typically 3 – 6 nm

Therefore, FRET efficiency declines very fast with larger distances. The radius contributes with the power of six to the E_{FRET} equation. The Foerster radius depends on the relative orientation of the fluorophores to each other. Nowadays, fluorophores with high quantum yields and high extinction coefficients are used. This leads to a Foerster radius between 4 – 6 nm. For the CFP / YFP combination, the Foerster radius is 5.2 nm[141]. Therefore, FRET can be measured at distances up to 10 nm; *ergo* the presence of FRET is a good indicator of close proximity of two proteins and implies biological meaningful protein-protein interactions. Thus, the combined method of FACS-based FRET is well suited to measure PPIs and study these in living cells.

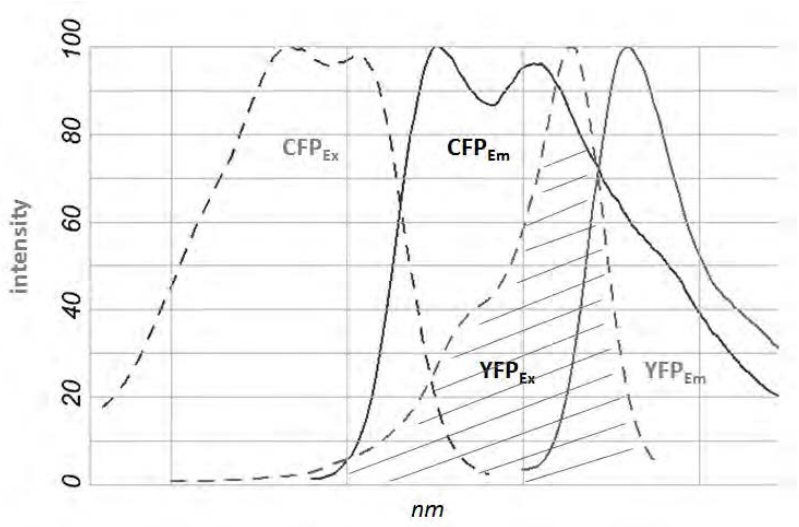

Fig. 4: **Overlap of CFP emission (CFP$_{Em}$) and YFP excitation (YFP$_{Ex}$) spectra.** Maxima of CFP excitation and emission are 405 nm and 450 nm, respectively. YFP shows an excitation maximum at 488 nm, an emission maximum at 529 nm.
Figure designed with the Fluorescence-SpectraViewer (Invitrogen).

1.1.6 *En Route* to the Three-Dimensional Structure of the HCV Glycoprotein E2

The object of the second part of this thesis was to establish an expression system for future three-dimensional structure determination of the two HCV envelope proteins E1 and E2 via X-ray analysis at atomic level using protein crystals.

E1 and E2 mediate virus-binding, entry into the host cell and induce membrane fusion. Hence, these two proteins mediate cellular infection and are crucial for the HCV replication cycle. Elucidating their three-dimensional structure would be a milestone for HCV research and therapy as well as vaccine development, since they are the main antigenic determinants of HCV. Earlier studies showed that

neutralizing antibodies against HCV E2 result in the protection against HCV infection[5]. Identifying the quaternary structure of HCV envelope proteins will contribute to vaccine development, complement existing data and give new implications for further research.

For such structural analysis, it is more feasible to use the secreted ectodomains of the membrane-proteins, since they can be isolated without any specific detergents, which eventually hamper the later crystallization. The amino acid sequences of the ectodomains can be defined, due to published pre-works. Different groups showed correct secretion and functionality. For instance Lorent et al.[143] determined the $E1_{326}$ ectodomain as an adequate truncated version. Regarding E2, Michalak et al.[144] indicated $E2_{661}$ as well suited for further analysis. Therefore, the $E1_{326}$ and $E2_{661}$ were used for expression.

It showed up, that at least the HCV envelope protein E2 is highly glycosylated with high-mannose N-glycans[6]. Protein glycosylation can play a crucial role regarding viral entry and adaptive immune response. In addition, glycosylation often influences protein folding and function. Therefore, mimicking the glycosylation-state can be crucial for HCV glycoproteins, and hence, another advantage of using secreted proteins passing ER and Golgi is that only fully glycosylated proteins will be harvested for further analysis. Thus, it is important to use an expression system, which is able to glycosylate the proteins in the same manner as in host cells and which can produce the protein in sufficient amount for further experiments.

By now, protein yield was the main drawback regarding glycoprotein expression[145]. In the present approach both ectodomains should be expressed within the Drosophila cell line S2, which provides a high-mannose glycosylation pattern, and yields high protein amounts as well.

2 Aims and Objectives

Since medical impact of HCV is very high and no vaccine is available, more insight into the viral life cycle is needed. Analogous to the study of Dimitrova et al.[61] on intra-viral PPIs of HCV non-structural proteins, the present work will extend the intra-viral PPI network implicating also structural proteins. Furthermore, this study will be performed in the context of living cells.

Prediction of HCV protein-protein interactions is not possible considering the lack of protein homologues in other viruses. However, virus PPIs are crucial since they mediate many mechanisms and functions, e.g. priming the host cell machinery for efficient replication. Within this context it is helpful to generate an intra-viral interactome also including short and dynamic interactions, which can easily be accomplished using FCET. For this, all ten HCV proteins should be generated as N- or C-terminal-tagged fusions and tested for their expression and localization to confirm their functionality. Extensive co-transfections in HEK293T cells should be done to establish the intra-HCV protein interaction network. Furthermore, it was planned to confirm the HEK293T HCV network in the liver cell line Huh7.5, which resembles a cell line supporting the full replication cycle of HCV. In addition, some of the novel interactions should be analyzed by biochemical approaches to assess if they also exert strong physical interaction within the cell.

In the second part of the present work, the goal was to establish an expression system for future elucidation of HCV E1 and E2 3D structure via X-ray crystallography and/or SAXS. The structural analysis of HCV proteins is crucial for vaccine development and the rational design of antivirals targeting the glycoprotein. Special epitope features on the glycoproteins can highlight for example new inhibitor docking sites. For the structural analysis it is essential that expressed proteins contain their specific post-translational modifications, since these modification are important for the proper folding of the respective protein. The proteins have to be expressed in high yields and need to be stable for longer periods. Additionally, their biological functionality has to be proven, which in return indicates correct folding. Optimization of protein purification shall result in high and pure yields, with little loss of protein to enable subsequent structural analysis.

3 Material

3.1 Nucleotides

3.1.1 Oligonucleotides

Name	Sequence	Target Gene
Primer for fluorescent fusion construction		
Core-XhoI f	atctcgagctagcacaaatcctaaacctc	Core
Core-EcoRI r	cagaattcttaagcagagaccggaacggtgatg	Core
E1-XhoI f	atctcgagctgcccaggtgaagaataccag	E1
E1-EcoRI r	cagaattcttacgcgtccaccccagcggccag	E1
E1-NheI f	catcggctagcatggcccaggtgaagaataccagtag	E1
E1-AgeI r	tcgaccggtgcacctgctcccgcgtccaccccagcggccagc	E1
E2-XhoI f	atctcgagctggcaccaccaccgttggagg	E2
E2-EcoRI r	cagaattcttatgcttcggcctggcccaac	E2
E2-NheI f	catcggctagcatgggcaccaccaccgttggagg	E2
E2-AgeI r	tcgaccggtgcacctgctcctgcttcggcctggcccaac	E2
p7-XhoI f	atctcgagctgcattggagaagttggtcg	p7
p7-EcoRI r	cagaattcttaggcataagcctgccggggc	p7
NS2-XhoI f	atctcgagcttatgacgcacctgtgcacgg	NS2
NS2-EcoRI r	cagaattcttaaaggagcttccacccttgg	NS2
NS3-XhoI f	atctcgagctgctcccatcactgcttatgc	NS3
NS3-EcoRI r	cagaattcttaggtcatgacctcaaggtcag	NS3
NS4A-XhoI f	atctcgagctagcacgtgggtcctagctgg	NS4A
NS4A-EcoRI r	cagaattcttagcattcctccatctcatc	NS4A
NS4B-XhoI f	atctcgagctgcctctagggcggctctcatc	NS4B
NS4B-EcoRI r	cagaattcttagcatgggatggggcagtcc	NS4B
NS5A-XhoI f	atctcgagcttccggatcctggctccgcg	NS5A
NS5A-EcoRI r	cagaattcttagcagcacacggtggtatcg	NS5A
NS5B-XhoI f	atctcgagcttccatgtcatactcctggac	NS5B
NS5B-EcoRI r	cagaattcttaccgagcggggagtaggaag	NS5B
Primer for non-tagged HCV proteins		
AgeI-ATG-Core	cataccggtatgagcacaaatcctaaacctc	Core
AgeI-ATG-E2	cataccggtatgggcaccaccaccgttgg	E2
AgeI-ATG-p7	cataccggtatggcattggagaagttggtc	p7
AgeI-ATG-NS2	cataccggtatgtatgacgcacctgtgcac	NS2
AgeI-ATG-NS3	cataccggtatggctcccatcactgc	NS3
Primer for ectodomain expression in S2 cells:		
E1e-BglII f	gcaagatctgcccaggtgaagaatac	E1
E1e-EcoRI r	gcagaattcgaagtaggccaagccgaac	E1
E2e-BglII f	gcaagatctggcaccaccaccgttg	E2
E2e-EcoRI r	gcagaattcgtcctccaagtcgcag	E2

3.1.2 DNA Ladder

Gene Ruler™ DNA Ladder Mix Fermentas GmbH (St. Leon-Rot) SM0333

3.1.3 PCR Nucleotides

dNTPs Stratagene (La Jolla, USA)

3.2 Plasmids

3.2.1 Fluorochrome Plasmids

From Clontech; Takara Bio Europe (Saint-Germain-en-Laye, France)

Name	Description
pECFP-C1	bearing the Cyan Fluorescence Protein; MCS downstream the fluorochrome
pEYFP-C1	bearing the Yellow Fluorescence Protein; MCS downstream the fluorochrome
pECFP-N1	bearing the Cyan Fluorescence Protein; MCS upstream the fluorochrome
pEYFP-N1	bearing the Yellow Fluorescence Protein; MCS upstream the fluorochrome

3.2.2 HCV Plasmid

Name	Description	Reference
pJFH1	wild type HCV from Japanese patient with fulminant Hepatitis C	[89]
pFK-luc-JFH1	reporter virus	[129]
pFK-luc-JC1	reporter virus	

3.2.3 Drosophila Expression System Plasmids

From the Invitrogen DES® kit, Life Technologies GmbH (Darmstadt)

Name	Description
pMT/V5/BiP-His	for inducible expression of secreted proteins
pMT/V5-GFP	GFP-control for positive transfection
pCoHygro	selection vector, encoding for Hygromycin resistance

3.3 Bacteria Strains

One Shot® Top10; chemically competent *Escherichia coli;* F- mcrA (mrr-hsdRMS-mcrBC) 80lacZM15 lacX74 recA1 ara139 (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG (Invitrogen, Karlsruhe/ Life technologies)

DH5-α: supE44, ΔlacU169, (ϕ80dlacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1

3.4 Eukaryotic Cell Lines

Name	Description
Huh7.5	human hepatoma cell line 7; clone 5
HEK293T	human embryonic kidney cell line, expression of large T antigen
Drosophila Schneider cells (S2)	derived from a primary culture of late stage (20 – 24 hours old) *Drosophila melanogaster* embryos (Invitrogen / Life Technologies, Darmstadt)

3.5 Media

3.5.1 Bacteria Media

The media / plates were dispensed and autoclaved (20 min, 120°C). After cooling to 50°C antibiotics were added.

Name	Components
Luria Bertani (LB) Medium	10 g/l bacto trypton 5 g/l yeast extract 8 g/l NaCl 1 g/l glucose The pH was set to 7.2 with NaOH Addition of 100 mg/l ampicillin or kanamycin
Luria Bertani (LB) agar plates	15 g agar ad 1 l LB media 1 mg/ml ampicillin or kanamycin
SOC medium	20 g/l bacto trypton 5 g/l yeast extract 2.5 mM NaCl 10 mM $MgCl_2$ 10 mM $MgSO_4$ 20 mM glucose

3.5.2 Cell Culture Media

Description	Components
Medium for Huh7.5 cell cultivation	DMEM (Dulbecco's modified Eagle Medium; Invitrogen/Gibco) with 350 µg/ml L-glutamine, 120 µg/ml streptomycin sulfate, 120 µg/ml penicillin und 10 % (v/v) heat inactivated FCS and 1 % (v/v) MEM *non essential amino acids* (NEAA)

Media for HEK293T cell cultivation	DMEM (Dulbecco's modified Eagle Medium; Invitrogen/Gibco) with 350 µg/ml L-glutamine, 120 µg/ml streptomycin sulfate, 120 µg/ml penicillin und 5 – 10 % (v/v) heat inactivated FCS
Media for S2 cell cultivation	Insect-XPRESS™ Media; Lonza (Basel, Switzerland)

3.5.3 Cell Culture Media Additives

Name	Company
L-Glutamine	PAA Laboratories GmbH (Cölbe)
NEAA	MEM non essential amino acids, PAA Laboratories GmbH (Cölbe)
Penicillin/Streptomycin	PAA Laboratories GmbH (Cölbe)
FCS (fetal calf serum)	Invitrogen/Gibco (Karlsruhe)
Hygromycin B	Merck (Nottingham, UK)
Metafectene Pro	Biontex (Martinsried)

3.6 Enzymes

3.6.1 Restriction Enzymes

New England ordered from Biolabs GmbH (Frankfurt) / Fermentas GmbH (St. Leon-Rot) used with the recommended buffers

3.6.2 Other Enzymes

Name	Company
0.05 % EDTA-**Trypsin**	Invitrogen/Gibco (Karlsruhe)
Alkaline Phosphatase	Roche (Mannheim)
T4-DNA-Ligase	Promega GmbH (Mannheim) & Roche (Mannheim)
Dream Taq™ DNA polymerase	Fermentas GmbH (St. Leon-Rot)
Pfu DNA polymerase	Fermentas GmbH (St. Leon-Rot)

3.7 Antibodies

3.7.1 Primary Antibodies

Antigen	Description, Dilution, Reference
Hepatitis C Virus antibodies	
HCV Core	monoclonal anti-mouse antibody (clone C7-50), detects AA21-40 of HCV Core, dilution: IF 1:50 & WB 1:500; Abcam (Cambridge, UK)

HCV E2	broad range neutralizing monoclonal anti-mouse antibody (clone AP33), IF 1:100, WB 1:100**;** Genentech, Inc. (San Francisco, USA; Owsianka *et al.*, 2005 ([11]))
HCV NS5A	monoclonal anti-mouse antibody (clone 2F6/G11) detects AA2054 to 2295 of HCV genome, dilution: IF 1:50; IBT (Reutlingen)
Other antibodies	
β-Actin	loading control for Western-blot, monoclonal anti-mouse antibody (clone AC-15), dilution WB 1:5,000, Sigma-Aldrich (Munich)
α-His	anti-mouse His-probe (H-3): sc-8036; Santa Cruz Biotechnology, Inc. (Heidelberg)
α-GFP	GFP monocl. anti-rabbit antibody; BioVision (California, USA)
α-GST	anti-Glutathione-S-Transferase antibody, IgG1 (mouse); clone G1, MCA1173 or clone vpg66 MCA1352; WB dilution 1:100; AbD Serotec, MorphoSys AbD GmbH, Düsseldorf

3.7.2 Secondary Antibodies

Antigen	Description, Dilution, Reference
α-rabbit	IRDye® 680 goat anti-rabbit IgG (H+L), Li-Cor Biotechnology GmbH (Bad Homburg) / IRDye® 700DX anti-rabbit IgG; Rockland antibodies & assays dilution WB 1:10,000
α-mouse	IRDye® 800 goat anti-mouse IgG (H+L), Li-Cor Biotechnology GmbH (Bad Homburg) / IRDye® 800 anti-rabbit IgG; Rockland antibodies & assays dilution WB 1:10,000

3.8 Chemicals

Name	Company
Agar	Carl Roth® GmbH & Co.KG (Karlsruhe)
Agarose	Carl Roth® GmbH & Co.KG (Karlsruhe)
Ampicillin	Ratiopharm GmbH (Ulm)
Beta-Mercaptoethanol	Merck KGaA (Darmstadt)
Bacto-Trypton	BD Biosciences Pharmingen (San Diego, USA)
Bromophenol blue	Merck KGaA (Darmstadt)
Chloroform	Carl Roth® GmbH & Co.KG (Karlsruhe)
cOmplete, ULTRA, Mini EDTA-free, EASYpack	Roche (Mannheim)
DEPC (Diethylpyrocarbonat)	Carl Roth® GmbH & Co.KG (Karlsruhe)

DTT (Dithiothreitol)	PAA (Cölbe)
Ethanol	Carl Roth® GmbH & Co.KG (Karlsruhe)
Ethidium bromide	Carl Roth® GmbH & Co.KG (Karlsruhe)
Ethylene diaminetetraacedic acid (EDTA)	Carl Roth® GmbH & Co.KG (Karlsruhe)
Glucose	Merck KGaA (Darmstadt)
Glutathione beads	GE healthcare
Glycine	PAA (Cölbe)
Glycerol	PAA (Cölbe)
HPLC water	AppliChem (Darmstadt)
Imidazole	Carl Roth® GmbH & Co.KG (Karlsruhe)
IPTG (Isopropyl-β-D-thiogalactopyranosid)	Carl Roth® GmbH & Co.KG (Karlsruhe)
Isopropyl alcohol	Carl Roth® GmbH & Co.KG (Karlsruhe)
Kanamycin	Ratiopharm GmbH (Ulm)
Lysozyme	AppliChem (Darmstadt)
Skimmed milk powder	Carl Roth® GmbH & Co.KG (Karlsruhe)
MES Pufferan 2-(N-Morpholino) ethansulfonsäure	Carl Roth® GmbH & Co.KG (Karlsruhe)
MTT Thiazolyl Blue Tetrazolium Bromide	Sigma-Aldrich, Chemie GmbH (Munich)
NaCl	PAA (Cölbe)
NaOH	Carl Roth® GmbH & Co.KG (Karlsruhe)
Paraformaldehyde (PFA)	Carl Roth® GmbH & Co.KG (Karlsruhe)
Phosphate Buffered Saline (PBS)	PAA (Cölbe)
Protein A sepharose	Sigma-Aldrich, Chemie GmbH (Munich)
Protein A	Sigma-Aldrich, Chemie GmbH (Munich)
Protein Assay Dye Reagent Concentrate	Bio-Rad Laboratories (Hercules, USA)
Sodium acetate	Promega GmbH (Mannheim)
Sodium acid	Carl Roth® GmbH & Co.KG (Karlsruhe)
Sodium dodecyl sulfate (SDS)	AppliChem (Darmstadt)
Tri-sodium phosphate dodecahydrat	Merck KGaA (Darmstadt)
Tris acetate EDTA Buffer (50X)	Bio-Rad Laboratories (Hercules, USA)
Triton X-100	AppliChem (Darmstadt)
UltraPure™ Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v⁄v)	Invitrogen (Darmstadt)
Yeast Extract	BD Biosciences Pharmingen (San Diego, USA)

3.9 Kits

Purpose	Name	Company
DNA isolation		
Mini prep	Resuspensionspuffer (P1)	Qiagen (Hilden)
	Lysepuffer (P2)	Qiagen (Hilden)
	Neutralisationspuffer (P3)	Qiagen (Hilden)
Midi prep	PureYield™ Plasmid Midiprep	Promega GmbH (Mannheim)

Maxi prep	Plasmid Plus Maxi Kit	Qiagen (Hilden)
DNA purification	Ultra Clean™ 15 DNA Purification Kit	Dianova GmbH (Hamburg)
DNA ligation	Takara DNA Ligationskit	Böhringer Ingelheim (Heidelberg)
RNA production	T7 RiboMAX™ Express Large Scale RNA Production System	Promega GmbH (Mannheim)
Virus detection	Luziferase Assay System	Promega GmbH (Mannheim)

3.10 Solvents and Buffer

Name	Components
FACS Buffer	2 % FCS; 1 mM EDTA in PBS
Mowiol	0.2 M Tris-HCl, pH 8.5; 12 % (w/v) Mowiol 4-88; 30 % (w/v) glycerin
DEPC Water	0.5 % (v/v) DEPC in H_2O_{dest} incubation o/n and autoclaved
Cytomix	120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4 (pH 7.6), 25 mM Hepes, 2 mM EGTA, 5 mM $MgCl_2$; pH adjusted to 7.6 with KOH; sterile filtration; **before electroporation addition of** 2 mM ATP (pH 7.6) and 5 mM Glutathione (pH 7.6)
10X HBS	8.18 g NaCl; 5.94 g HEPES; 0.25 g Na_2HPO_4 x $2H_2O$; ad 100 ml H_2O_{dest}, store at -20°C; dilute to 2X HBS and adjust pH to 7.23; store at -20°C

3.11 Western-blot and Proteins

Name	Components
Ponceau-S	0.1 % (w/v) Ponceau S; in 5 % acetic acid
Coomassie	0.1 % Coomassie R250, 10 % acetic acid, 40 % methanol
1X TGS	30.3 g glycine, 150.14 g Tris, ad 10 l H_2O_{dest}
2X Towbin	60.5 g Tris, 300.3 g glycine, 10 g SDS, 4 l MeOH, ad 10 l H_2O_{dest}
20X TBST	1 M Tris, 18 % (w/v) NaCl, pH adjusted to 7.6, addition of 1 % (v/v) Tween20
Block Buffer	10 % powdered skimmed milk in 1X TBST

RIPA Buffer	150 mM NaCl, 50 mM Tris-HCl, 1 % Nonidet P-40, 0.5 % sodium desoxycholate, 0.1 % Na-SDS, 5 mM EDTA
Laemmli Buffer	4 ml SDS 10 %, 1 ml β-Mercapto-EtOH, 2 ml glycerin, 0.2 ml EDTA (1 M), 1 ml Bromophenol blue 0.1 %, 1.3 ml Tris (1 M); pH 6.8, 0.5 ml H_2O_{dest}
Protein Ladder	Spectra™ Multicolor Broad Range Protein Ladder Fermentas GmbH (St. Leon-Rot)
Glutathione Sepharose 4B	GE Healthcare (Munich)
Ni-NTA Sepharose	Thermo Scientific (Bonn)
Ni-NTA Column – HisTrap FF 1ml	GE Healthcare (Munich)
Ni-NTA Column – Excel 1ml	GE Healthcare (Munich)
Superdex 200 Column (Gel filtration)	GE Healthcare (Munich)
UNO™ S-6 (6ml) (Ion Exchange)	Bio-Rad (Munich)

3.12 Consumables

Name	Company
Amicon Ultra-0.5; 4; 15	Millipore (Schwalbach)
Cryo-Tubes	Sarstedt (Nümbrecht)
MILLEX GP 0.2 µm filter	Millipore (Schwalbach)
MILLEX GP 0.45 µm filter	Millipore (Schwalbach)
Nitrocellulose membrane 0.4 µm	Schleicher & Schuell (Dassel)llg
Pipette tips (10 - 1000 µL)	Sarstedt (Nümbrecht)
Pulser cuvette, 0.4 cm	Bio-Rad Laboratories (Hercules, USA)
Tissue culture flasks (T25, T75, T175)	Sarstedt (Nümbrecht)
Tissue culture plates	Greiner bio-one (Frickenhausen)
Tubes (0.5 ml, 1.5 ml, 2 ml)	Sarstedt (Nümbrecht)
Tubes (15 ml, 50 ml)	Sarstedt (Nümbrecht)
Vivaspin 500 µl / 15 ml	Sartorius Stedim Biotech (Göttingen)
Whatman paper	GE Healthcare (Munich)

3.13 Equipment

Name	Company
DNA gel electrophorese system	Bio-Rad Laboratories (Hercules, USA)
Eppendorf centrifuge 5417 R	Eppendorf (Hamburg)
Eppendorf centrifuge 5810 R	Eppendorf (Hamburg)
Eppendorf Multipette® plus	Eppendorf (Hamburg)
BD FACSCantoII™	Becton Dickinson, Immuncytometry Systems, (San José, USA)

Gene Pulser Xcell System Electroporated	Bio-Rad Laboratories (Hercules, USA)
GeneAmp® PCR System 9700	AB Applied Biosystems (Darmstadt)
Gilson Pipetman®	Gilson Inc. (Middleton, USA)
HERAsafe® Incubator	Thermo Fisher Scientific GmbH (Hanau)
HERAsafe® sterile bench	Thermo Fisher Scientific GmbH (Hanau)
Infinite® M200 plate reader	Tecan Group Ltd. (Männedorf, Schweiz)
Nanodrop	PEQLAB Biotechnologie GmbH (Erlangen)
Nikon Eclipse Ti	Nikon (Tokyo, Japan)
Odyssey Imaging System	Li-Cor Biotechnology GmbH (Bad Homburg)
Pipettboy acu IBS INTEGRA	Biosciences (Chur, Schweiz)
Pipettes (5 ml, 10 ml, 25 ml)	Sarstedt (Nümbrecht)
Shaker Innova® 43 New	Brunswick (Nürtingen)
Synergy plate reader	BioTek, (Bad Friedrichshall)
Thermoblock	Eppendorf (Hamburg)
UV-Transilluminator GelDoc 2000	Hartenstein (Würzburg)
Vortex-Genie 2	Scientific Industries (New York, USA)
Zeiss 510 Meta	Carl Zeiss (Jena)

3.14 Software and Databases

Acrobat X	PDF data processing	Adobe
Cytoscape 2.8.2	Network Analysis	Open Source; www.cytoscape.org
FACSDiva	Control Software and Data Processing for Flow Cytometers	BD biosciences
Gen5™	Microplate Reader Software	Synergy™; BioTek
i-control™	Microplate Reader Software	TECAN
LSM Image Browser	Management of CLSM images	Zeiss
LSM510 Software	Control Software for Zeiss LSM	Zeiss
Microsoft Office 2010	Text processing	Microsoft
Microsoft Office 2013 (Mac)	Text processing	Microsoft
Odyssey® Software	Scanning Western-Blots	Li-Cor
Papers 2.4.6	Reference organization	Mekentosj
Photoshop CS4	Image processing	Adobe
Prism 5 for Mac OSX	Statistical analysis	GraphPad
PubMed	Literature database	Open Software (NCBI)
Serial Cloner 2.5	Sequence data processing	Serial Basics
UNICORN 4.12	Control Software for ÄKTA™	GE Healthcare Life Sciences
VirusMINT	Viral protein interaction db	mint.bio.uniroma2.it

4 Methods

4.1 Molecular-Biological Methods

4.1.1 Cultivation of Bacteria (for Plasmid Isolation)

All used plasmids contained either a kanamycin or an ampicillin resistance cassette. Therefore, transformed bacteria were selected with 100 µg/ml ampicillin and 50 µg/ml kanamycin respectively, in culture medium. The used *E. coli* strain One Shot® Top10 (Invitrogen) and DH5α were cultivated at 37°C on LB agar plates or in LB medium.

4.1.2 Isolation of Plasmid DNA

Isolation of plasmid DNA was performed via alkaline bacteria lysis[146] followed by purification. According to the amounts of DNA needed, the DNA extraction was done via a Mini, Midi or Maxi preparation following the manufacturer's instructions of the Qiagen and Promega kits (see materials). DNA preps were solved in DNAse free water, concentration and quality control was confirmed via NanoDrop, which measures the DNA absorption maximum at 260 nm and via electrophoretic DNA separation on an 0.5 to 1.5 % agarose gel, to control DNA fragment sizes.

4.1.3 Polymerase Chain Reaction – PCR

Polymerase Chain Reaction (PCR), which was established by Kary Mullis[147], enables amplification of specific DNA fragments in vitro between two oligonucleotide primers in a cyclic process from denaturation over primer hybridization to elongation. In this case, a PCR thermo cycler of Applied Biosystems was used with the standard PCR conditions:

96°C	5 min	initial denaturation / first step delay	
96°C	30 sec	denaturation	for 35 cycles
52°C	45 sec	annealing / hybridization of the oligonucleotides	
72°C	2 min	elongation	
72°C	8 min	finale elongation / last step delay	
4°C	∞		

Depending on primers and templates, the conditions for the thermocycler vary. The standard-PCR reaction mix for one sample (50 µl prep) was:

10X polymerase buffer	5.0 µl
dNTPs 10 mM	0.5 µl
Primer I (200 µM)	0.75 µl
Primer II (200 µM)	0.75 µl
Taq-Polymerase (5U/µl)	0.25 µl
Template DNA	0.5 µl (100 to 500 ng DNA)
H_2O_{dest}	40 µl

Control of PCR fragments was performed via agarose gel.

4.1.4 DNA Purification from Agarose Gels and Solutions

For the DNA extraction out of an agarose gel the Ultra Clean™ 15 DNA Purification Kit of Dianova was used following the manufacturer's instructions.

4.1.5 Restriction Digest of DNA and PCR Fragments

Restriction endonucleases are part of the restriction/modification system that protects bacteria from the uptake of foreign DNA and assure genetic variability. The used type II restriction endonucleases allow specific cutting/editing of dsDNA. Dependent on the enzyme, different conditions regarding temperature, buffers and incubation times are needed. The standard mixture was:

Plasmid DNA	2 µl
10X Buffer	2 µl
Restriction enzyme I	0.5 µl
Restriction enzyme II	0.5 µl
ad 20 µl H_2O_{dest}	

The mixture was incubated at 37°C for 30 min up to 2 hours followed by a restriction control via agarose gel electrophorese.

4.1.6 DNA Fragment Ligation

Recombinant DNA molecules are often produced by insertion of a specific DNA fragment (insert) in a cloning vector with the help of a T4 ligase, which ligates under the use of ATP, free 3' OH ends to 5' phosphate ends of dsDNA. It is possible to link sticky and blunt ends. The insertion into the vector requires same restriction sites in vector and insert, which were introduced via the primers. For this approach, ligation reaction was performed with the Takara DNA ligation kit (Boehringer Ingelheim) where plasmid DNA and insert were added to the T4 DNA ligase with a ratio of 1:4 according to the manufacturer's instructions:

Vektor	0.5 µl
Insert	2 µl
Solution I (T4 ligase + buffer)	2.5 µl

The mixture was incubated 3 hours at 16°C followed by an overnight incubation at 4°C.

4.1.7 *E. coli* One Shot® Top10 Transformation

To 10 µl of bacteria, which were thawed on ice, 1 µl ligation preparation or plasmid DNA was added, respectively followed by 15 minutes incubation on ice. Bacteria were heat shocked at 42°C for 50 seconds and incubated on ice for 2 minutes according the addition of 150 µl SOC medium and incubation for 60 min at 37°C, while shaking. The transformed bacteria were plated on agar plates with the corresponding antibiotic and incubated overnight at 37°C. For analyzing the correct insertion, a control restriction with specific restriction enzymes was performed followed by separation on an agarose gel.

4.1.8 DNA Sequencing

Sequencing was done commercially via Eurofins MWG GmbH (Ebersberg) and Seqlab (Sequence Laboratories, Göttingen) according to their protocols.

4.1.9 Glycerin Stocks

For longer storage periods of transformed bacteria, an overnight culture was diluted 1:1 with a 50 % glycerin and 2.9 % NaCl solution in a total volume of 1.5 ml and stored at -80°C.

4.2 Cell-Biological Methods

4.2.1 Freeze and Thawing of Eukaryotic Cells

Cells were thawed in a 37°C water bath, diluted in 50 ml pre-warmed medium and centrifuged to wash out DMSO (from the freezing medium). 5 to 10 ml fresh medium was added followed by cell seeding.

For freezing, cells of confluent flasks were detached, washed with PBS and suspended in a $5x10^6$ cells/ml precooled freezing suspension. After 24h at -80°C the cells were transferred to liquid nitrogen for longer storage periods. To avoid formation of ice crystals in and out of the cytoplasm

and to bypass dehydration 10 % DMSO (dimethyl sulfoxide) was added to the freezing medium (10 % (v/v) DMSO, 50 % medium, 40 % (v/v) FCS).

4.2.2 Sub-Cultivation of Eukaryotic Cells

Incubation of cells was performed at 37°C, 5 % CO_2 and a humidified atmosphere (95 % relative humidity). Cells were cultivated in DMEM (Dulbecco's modified Eagle Medium, Biochrom) with these additives: L-glutamine [1 % (v/v)], streptomycin sulphate & penicillin [1 % (v/v)] and heat inactivated FCS [10 % (v/v)]. Determination of viable cell count was done with a Neubauer counting chamber.

4.2.2.1 Cultivation of Huh7.5 Cells

The human liver hepatoma cell line Huh7.5 needs 1 % (v/v) MEM non-essential amino acids (NEAA) as additional medium supplement. The adherent cells were passaged at a confluence of 80 %. For this, cells were washed, trypsinized and suspended in FCS containing medium to inactivate the trypsin-reaction. Cells were split in a 1:4 to 1:10 ratio and seeded.

4.2.2.2 Cultivation of HEK293T Cells

The human embryonic kidney cells are a semi-adherent cell line and do not require trypzinization. Gentle rocking of the tissue flask is sufficient to loose cells from the bottom. These cells were passaged in a 1:10 ratio nearly every two days when they reached confluence.

4.2.2.3 Cultivation of Drosophila Schneider (S2) Cells

S2 cells, derived from a primary culture of late stage *Drosophila melanogaster* embryos do not require CO_2 and grow at room temperature at a loose, semi-adherent monolayer in tissue culture flasks, in suspension in spinner and shake flasks. These cells were cultivated with Lonza Insect-XPRESS^{TM} protein-free insect cell medium with L-glutamine. As additives streptomycin sulphate & penicillin [0.5 % (v/v)] were added with or without 10 % FCS.

4.2.3 Transfection of Eukaryotic Cells

4.2.3.1 Transfection of HEK293T Cells via Calcium Phosphate

The human embryonic kidney cells were transfected via the calcium phosphate method, which was described in 1973 by Frank L. Graham[148]. In this approach, a DNA-phosphate complex forms by mixing DNA, 2 M $CaCl_2$ and 2X HBS in specific ratios. The complex reaches the cell presumably via endocytosis[149]. Analysis of the cells was performed 24h after transfection.

4.2.3.2 Transfection of Huh7.5 Cells with Metafectene Pro

The liver cell line was transfected with Metafectene Pro from Biontex according to the manufacturer's protocol. Analysis of the transfected cells was performed 48h after transfection.

4.2.4 Confocal Analyses – Localization and Co-Localization Studies in Kidney and Liver Cells

For the microscopic analyses with a Zeiss LSM510 Meta, HEK293T cells were seeded on 12 mm coverslips in 6-well plates with 400.000 cells/well, Huh7.5 cells in 12-well plates with 350.000 cells/well respectively, approximately 12 hours pre transfection with calcium phosphate or Metafectene® Pro. 24h/48h later the cells were washed with PBS and fixed with 2 % PFA for 20 to 30 minutes at 4°C. Cells were washed and mounted with Mowiol on microscope slides. Mowiol has the same refractive index as immersion oil. Conversion of images was achieved with the Zeiss LSM Image Browser and Photoshop.

4.2.5 FACS-Based FRET

Cloning strategy was realized as described before[90]. To get YFP[a] and CFP[b] fusions Clontech cloning vectors were used (kind gifts from Dr. Klaudia Giehl, University of Ulm). Due to C-terminal cleavage sites of HCV proteins, all fusions were N-terminal tagged. Each HCV protein was constructed as CFP and as YFP fusion. Therefore, every interaction pair could be tested in two combinations (CFP-protein A with YFP-protein B and YFP-protein A witch CFP-protein-B).

[a] EYFP and YFP are used equivalent

[b] ECFP and CFP are used equivalent

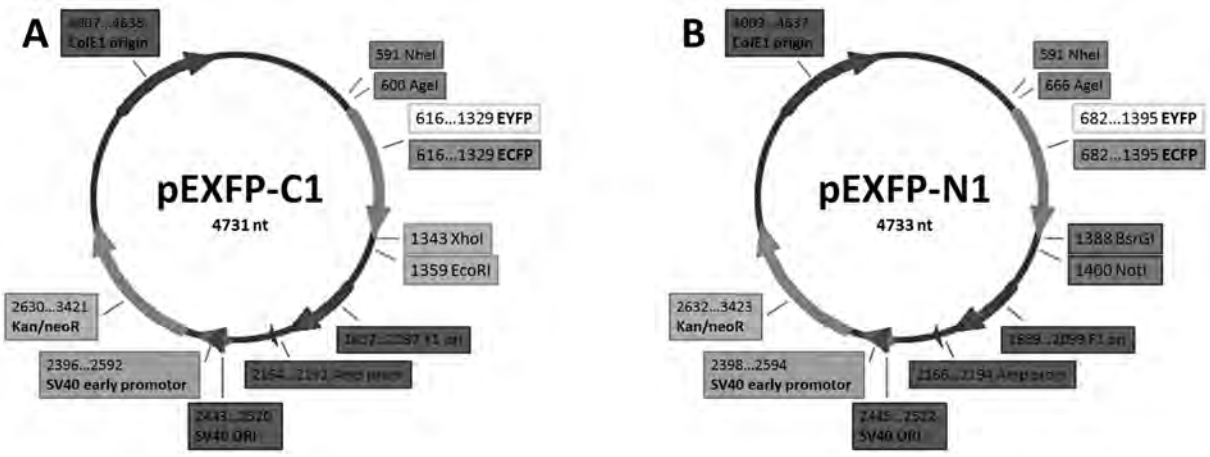

Fig. 5: **Expression vectors for HCV fusion constructs.** (A) pEYFP-C1 and pECFP-C1 (indicated as pEXFP-C1), including the multiple cloning site (MCS) downstream the fluorochrome coding sequence, with XhoI and EcoRI restriction sites to design N-tagged HCV fluorochrome fusions. C1-vectors were used for the most fusions. (B) pEYFP-N1 and pECFP-N1 (indicated as pEXFP-N1), containing an upstream MCS, and additional BsrGI and NotI restriction sites located downstream of XFP, for the construction of E1, NS3 and NS5A fusions, due to internal XhoI restriction sites in their coding sequences.

Each HCV protein was constructed, either as CFP-fusion, acting as energy donor, or as YFP-fusion, responding as energy acceptor. All ten HCV proteins were amplified from the HCV JFH1 sequence (Uniprot Q99IB8) and ligated into the pEXFP-C1 (pECFP-C1 & pEYFP-C1) or pEXFP-N1 (pECFP-N1 & pEYFP-N1) vectors (Clontech) depending on the presence of internal restriction sites (HCV E1, NS3 and NS5A all contain an internal XhoI site; see also Fig. 6). Primer sequences can be found in the material section of this thesis. All HCV proteins were N-terminally tagged, due to the presence of C-terminal protease cleavage sites. The non-tagged constructs were amplified with the same reverse primers as the XFP-fusions; nevertheless new forward primers were designed with an additional ATG and NheI restriction site, inserted into the pEYFP-C1 vector replacing the YFP sequence. Identity of the constructs was confirmed by commercial sequencing (Seqlab, Göttingen or MWG, Ebersberg).

Fusion Construct	bp	kDa Literature		kDa +XFP
pE**Y**FP-C1 HCV **Core** XhoI EcoRI	573	20 [45]	21 [43, 44, 46]	**52**
pE**C**FP-C1 HCV **Core** XhoI EcoRI				
pE**Y**FP-**N1** HCV **E1** BsrGI NotI	576	21 [43, 45]	30-35 [44, 53]	**66**
pE**C**FP-**N1** HCV **E1** BsrGI NotI				
pE**Y**FP-C1 HCV **E2** XhoI EcoRI	1101	40 [43, 45]	70 [44, 53]	**71**
pE**C**FP-C1 HCV **E2** XhoI EcoRI				
pE**Y**FP-C1 HCV **p7** XhoI EcoRI	189	7 [44]		**38**
pE**C**FP-C1 HCV **p7** XhoI EcoRI				
pE**Y**FP-C1 HCV **NS2-3** XhoI EcoRI	651	23 [44, 53]	24 [43]	**55**
pE**C**FP-C1 HCV **NS2-3** XhoI EcoRI				
pE**Y**FP-**N1** HCV **NS3** BsrGI NotI	1893	68 [43]	70 [44, 53]	**101**
pE**C**FP-**N1** HCV **NS3** BsrGI NotI				
pE**Y**FP-C1 HCV **NS4A** XhoI EcoRI	162	6 [43]	8 [44, 53]	**39**
pE**C**FP-C1 HCV **NS4A** XhoI EcoRI				
pE**Y**FP-C1 HCV **NS4B** XhoI EcoRI	783	27 [44, 53]	29 [43]	**60**
pE**C**FP-C1 HCV **NS4B** XhoI EcoRI				
pE**Y**FP-**N1** HCV **NS5A** BsrGI NotI	1398	29 [43]	56 + 58 [44, 53, 37]	**89**
pE**C**FP-**N1** HCV **NS5A** BsrGI NotI				
pE**Y**FP-C1 HCV **NS5B** XhoI EcoRI	1773	64 [43]	68 [44, 53]	**99**
pE**C**FP-C1 HCV **NS5B** XhoI EcoRI				

Fig. 6: **Overview of generated HCV JFH1 proteins fused to YFP and CFP.** Listed are all 20 constructs and the respective vector, which was used for ligation, the insert size in basepairs and the expected protein weight in kDa with and without XFP-tag.

HEK293T and Huh7.5 cells were transfected with the fusions. For the FACS analysis a FACSAria Cytometer (BD Bioscience) equipped with 405 nm, 488 nm and 633 nm lasers was used[90]. In brief, the detection of YFP, CFP and FRET signals were performed separately (Fig. 7). For YFP, signals cells were exited with 488 nm and the resulting signal detected with a 529/24 filter (Semrock). CFP signals were determined via the 450/40 filter after excitation at 405 nm. To measure FRET signal, cells were exited with 405nm followed by signal detection with the 529/24 filter (Semrock) again.

Five specific controls for each co-transfection setup were used (Fig. 7; A, B, C). Mock cells were transfected with water; the vectors pECFP-C1 and pEYFP-C1 were transfected to exclude false positive FRET signals (Fig. 7, B; setting the P2 gate). Co-transfection of pECFP and pEYFP served as FRET negative control, again to exclude false positive FRET signals and background (Fig. 7, C; setting the P3 gate). In the end, a construct where CFP is fused to YFP served as positive control (CFP--YFP). A minimum of 3,000 CFP and YFP co-transfected cells was analyzed. Data analysis was performed with BD Bioscience FACSDiva™ Software.

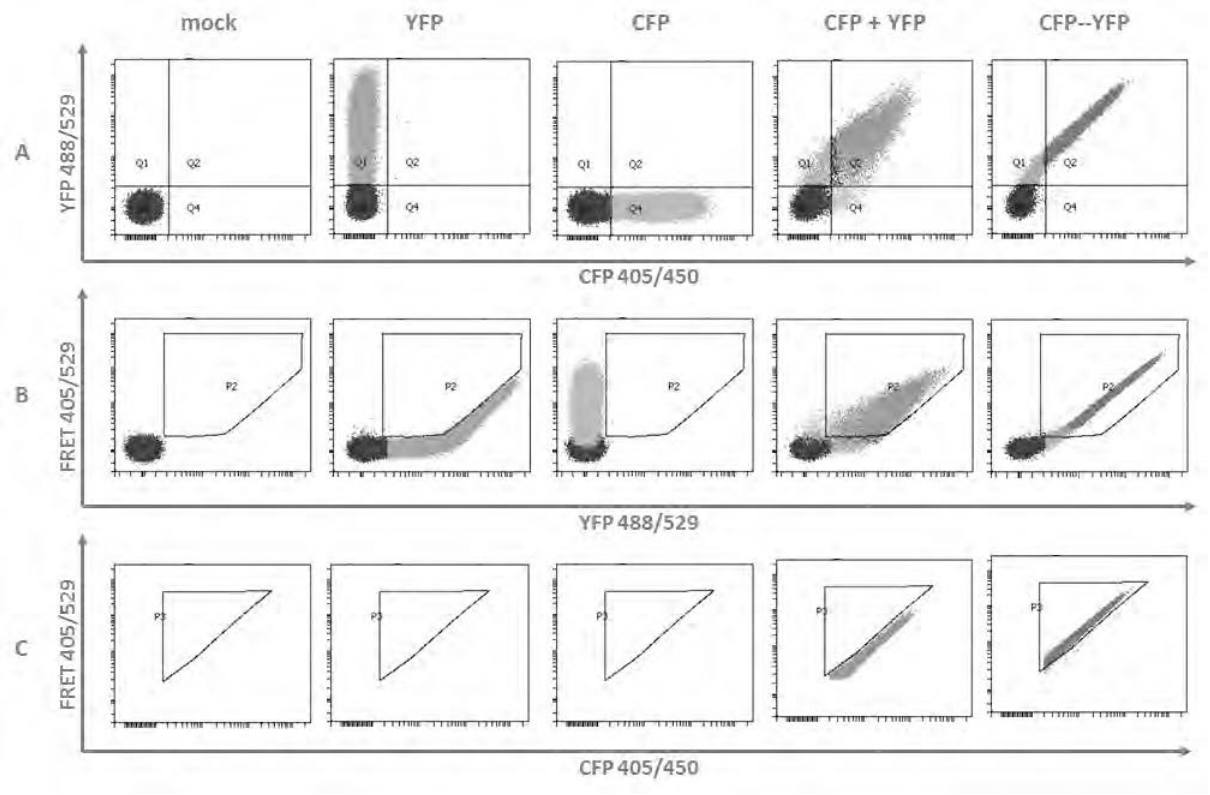

Fig. 7: **Controls and gating strategy for the FACS-based FRET approach.** For each co-transfection of HEK293T and Huh7.5 cells, five distinct controls were used to adjust the gates for measurement with a flow cytometer properly. Cells were transfected with water (mock), YFP, CFP, YFP and CFP as FRET negative control and CFP—YFP (CFP fused to YFP) as FRET-positive control. YFP signals against CFP signals are depicted in A, to detect single and co-transfection rates. Co-transfected cells were analyzed further, indicated in panel B, where FRET signals against YFP signals are indicated. False FRET-positives arise while exciting YFP at 405 nm. These false positives have to be excluded with the gate indicated in B (P2). Cells from this gate were further analyzed for FRET signals, indicated in panel C, where FRET signals are depicted against CFP signals. The (P3) gate was altered to differentiate between FRET-negative (YFP + CFP co-transfected) and FRET-positive (CFP--YFP transfected) cells.

4.2.6 RNA Production & Electroporation of Huh7.5 Cells

The HCV vectors (see 3.2.2) contain a T7 promotor upstream the HCV genome. After linearization an *in vitro* transcription of RNA with the Promega T7 RiboMAX™ Express Large Scale RNA Production System including a T7 RNA polymerase was performed.

After quality control via NanoDrop and electrophoresis, RNA was transfected into the permissive cell line Huh7.5 via electroporation (Gene Pulser Xcell System Electroporator from Bio-Rad), according to the Nature protocol from Kato et al.[150] with some alterations. Cells for electroporation were washed, trypsinized and suspended. For each electroporation $6x10^6$ cells were centrifuged at 700 rpm for five minutes, washed with PBS, centrifuged again and suspended in 400 µl Cytomix with freshly added (end concentration) 2 mM ATP and 5 mM glutathione. After transfer into electroporation cuvettes, 5 to 15 µg RNA (thawed on ice) were added to this mixture. The pulse was given under the conditions of 975 µF and 270 V, which should not take longer than 25 milliseconds. Cells were transferred directly into prepared 50 ml falcons with 20 ml complete medium and seeded into well plates or cell tissue flasks. To remove dead cells, medium was changed four to eight hours after electroporation.

4.2.7 Harvesting Virus Containing Supernatant & Infection

48 to 72 hours later, virus particle containing supernatants of electroporated cells were harvested and filtered through a 0.45 µm pore filter and stored at -80°C. Huh7.5 cells were seeded at least 4h pre infection. Their medium was then exchanged via virus containing supernatant. Medium was refreshed after eight hours incubation and analysis of infected cells was performed 48 to 72 hours later.

4.2.8 Luciferase Assay

Promega Luciferase Assay was performed by following the manufacturer's recommendations, which permits sensitive and rapid quantitation of reporter virus RNA from electroporated Huh7.5 cells. Firefly luciferase converts luciferin via oxidation by utilizing $ATP \bullet Mg^{2+}$ as co-substrate to oxyluciferin, resulting in luminescence. The level of luminescence is directly dependent on the amount of luciferase expression. The used pFK-luc vectors encode a Firefly luciferase from *Photinus pyralis* as reporter gene upstream of the HCV genes. Electroporated cells were lysed with 30 µl lysis buffer for 30 min at 4°C. Directly before measurement of luciferase intensity, 20 µl of lysed cells were transferred onto a white, flat bottom 96-well plate with the addition of 40 µl luciferin.

4.2.9 MTT Viability Assay

Measurement of living cells was performed via a colorimetric MTT assay (MTT: 3-(4, 5-Dimethylthiazol-2-yl)-2, 5-diphenyltetrazoliumbromid[151] to determine the linear relationship between absorbance and cell concentration. MTT is a pale yellow substrate that is cleaved by living cells via the NADH-dependent succinate dehydrogenase, to yield a dark blue insoluble formazan product. This process requires active mitochondria. Even freshly dead cells do not cleave significant amounts of MTT. For the assay, cells were seeded in a 96-well plate in 100 µl medium and treated with substances to test their influence on cell viability. Eight hours after incubation (or o/n) with the substrates, 10 µl MTT were added to the cells for further three to four hours incubation at 37°C. Exchange of the medium with 100 µl ethanol/DMSO (1:1) dissolves formed formazan crystals while shaking for 10 to 20 minutes. Absorption was measured with a plate reader at 570 nm with a reference wavelength at 630 nm. Cell survival rate was calculated via comparison to non-treated cells.

4.3 Biochemical Methods

4.3.1 Cell Lysis

For the analysis of whole cell protein lysates, cells were harvested and washed. The centrifuged pellet can be stored at -20°C or it can be proceeded with cell lysis. For lysis the pellet was suspended in 400 µl NP-40 or Ripa buffer. After incubation for 30 minutes on ice, cells were sonicated two times for 30 pulses (pulse duration: one second; output control: 8; duty cycle: 80 %). Cell debris was pelleted via centrifugation for 10 min at full speed and 4°C; supernatant could be stored for longer periods at -80°C.

4.3.2 Bradford-Assay

Measurement of protein concentration was performed with the Bio-Rad Protein Assay by following the manufacturer's recommendations. 800 µl PBS and 200 µl reagent were provided in a cuvette and 1 µl of protein lysate was added, mixed and incubated for five minutes at room temperature. Additional standard dilutions of BSA and a blank were prepared. Concentration measurement was performed at 595 nm. This assay has a sensitivity of 200 - 1500 µg protein per ml.

4.3.3 Discontinuous SDS-Polyacrylamide Gel Electrophorese (SDS-PAGE)

SDS-PAGE enables the discrete separation of proteins according to their size. The electrophorese principles were developed by Arne Tiselius in the early 20th century[152].

After polymerization of the separation gel, the mixture for the sample gel can be added. Due to the different pH in sample and separation gel, protein samples are first compressed into a thin starting band and finally resolved and separated. Casted gels were placed onto gel supports (Bio-Rad), upper and lower chambers were filled with TGS; 2X or 5X Laemmli buffer was added to the prepared protein lysates, boiled at 95°C for 5 minutes, and loaded onto the gel with a protein standard. Electrophoretic separation took place at 20 mA per gel for about two hours in TGS buffer. The next step was Coomassie staining or Western-blotting.

Recipe for casting a gel:

Separation gel:	10 % (10ml)	12 % (10ml)
		ml
H_2O	4	3.3
30 % acryl-bisacrylamide mix	3.3	4
1.5 M Tris (pH 8.8)	2.5	2.5
10 % SDS	0.1	0.1
10 % ammonium persulfate	0.1	0.1
TEMED	0.016	0.016

Sample gel:	10 ml
	ml
H_2O	6.95
30 % acryl-bisacrylamide mix	1.7
1 M Tris (pH 6.8)	1.25
10 % SDS	0.1
10 % ammonium persulfate	0.1
TEMED	0.01

4.3.4 Western-Blot

Transfer of the separated proteins from the gel to a nitrocellulose membrane (0.4 µm; Schleicher und Schuell) was performed via tank blotting (BioRad) and Towbin buffer. The transfer could be proved with Ponceau-S. Blocking of the membrane with 10 % skimmed milk in TBST or PBST for 1 hour avoided unspecific antibody binding. The primary and secondary antibodies were diluted in TBST [sodium acid (0.1 %) was added to primary antibody for storage and reuse]. The membrane was incubated with primary antibodies for at least two hours or overnight. To remove unspecific bound antibodies, the membranes were washed three times with TBST or PBST, each for 10 minutes, followed by incubation with the IRDye conjugated secondary antibody for 20 – 40 minutes. Again, the membrane was washed three times before detection of protein bands with Odyssey infrared imaging system (LI-COR).

4.3.5 Ponceau-S Staining

Western-blot membranes were stained with Ponceau S solution for maximum five minutes. To reduce background, the bound solution was washed with distilled water from the membrane. A minimum of 5.0 µg protein per band can be detected with this assay.

4.3.6 Coomassie Staining

SDS gels were stained with Coomassie solution for at least two hours. De-staining was performed overnight with distilled water, until protein-bands distinguish from background. A minimum of 0.5 µg protein per band can be detected with this assay.

4.3.7 Co-Immunoprecipitation (Co-IP)

All steps were performed on ice or 4°C, to avoid proteolysis with the addition of protease inhibitors to the NP-40 buffer. After lysis of transfected or electroporated cells with 600 µl NP40 buffer for 30 minutes, the cells were sonicated twice for 30 pulses (pulse duration: one second; output control: 8; duty cycle: 80 %). Protein concentration was determined via Bradford assay. For each sample, 2 mg protein-A sepharose were needed, which had to be swelled for one hour by nutating in 500 µl NP-40 buffer. The relevant antibody was given to the beads in an appropriate amount (about 0.5 µl per sample, depending on the antibody), incubation and agitation for one to two hours. 3 mg of each sample lysate was adjusted to 500 µl. To avoid unspecific binding, 50 µl pansorbin were added followed by incubation for one hour under agitation. Antibody coupled sepharose and lysates were centrifuged for five minutes at 6000 rpm, followed by three times washing with NP-40. The supernatants of the lysates were incubated and nutated with the aspirated sepharose overnight. After three times washing, 15 µl of 2X Laemmli buffer was added and boiled at 95°C for five minutes and centrifuged for further 10 minutes at 14,000 rpm and 10°C. Samples were analyzed via SDS PAGE and Western-blot.

4.4 HCV Ectodomain Expression

HCV glycoprotein ectodomains were expressed for further structural analyses. Transmembrane domains, which anchor the proteins in the ER membrane, were deleted. Via the passage over ER and Golgi, only completely post-translational modified proteins were secreted into the supernatant, which is crucial for correct structural analyses as well.

Various studies showed different properties for truncated forms of HCV glycoproteins, only varying in the length of deletion. Functional truncated forms for E1 (AA 192 – 326) were shown by Eric Lorent et al.[143] Mar Rodríguez-Rodríguez et al.[7] and Thomas Krey et al.[153] presented a functional ectodomain for E2 (AA 384 – 661). Thus these sequences were used in the present work for prospective structural analyses.

The two HCV envelope proteins are highly glycosylated. E1 bears up to six glycosylation sites (five strongly conserved) and E2 11 glycosylation sites (nine are strongly conserved). This glycosylation can play a crucial role for protein folding, entry function and modulating the immune response[13]. All 11 E2 glycosylation sites were proved to be occupied by high mannose glycans[6].

4.4.1 Amino Acid Sequences of E1 and E2 Ectodomains:

E1e
A Q V K N T S S S Y Met V T N D C S N D S I T W Q L E A A V L H V P G C V P C E R V G N T S R C W V P V S P N Met A V R Q P G A L T Q G L R T H I D Met V V Met S A T F C S A L Y V G D L C G G V Met L A A Q V F I V S P Q Y H W F V Q E C N C S I Y P G T I T G H R Met A W D Met Met Met N W S P T A T Met I L A Y V Met R V P E V I I D I V S G A H W G V Met F G L A Y F

E2e
G T T T V G G A V A R S T N V I A G V F S H G P Q Q N I Q L I N T N G S W H I N R T A L N C N D S L N T G F L A A L F Y T N R F N S S G C P G R L S A C R N I E A F R I G W G T L Q Y E D N V T N P E D Met R P Y C W H Y P P K P C G V V P A R S V C G P V Y C F T P S P V V V G T T D R R G V P T Y T W G E N E T D V F L L N S T R P P Q G S W F G C T W Met N S T G F T K T C G A P P C R T R A D F N A S T D L L C P T D C F R K H P D A T Y I K C G S G P W L T P K C L V H Y P Y R L W H Y P C T V N F T I F K I R Met Y V G G V E H R L T A A C N F T R G D R C D L E D

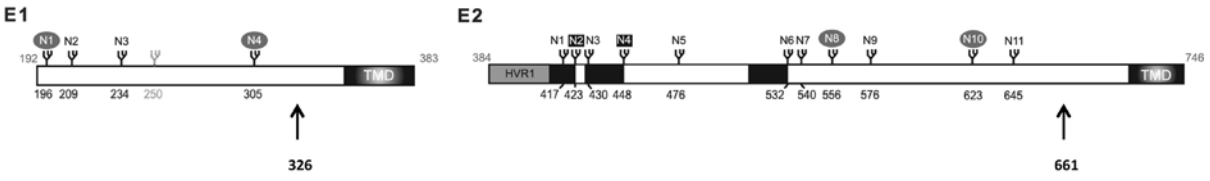

Fig. 8: **Glycosylation sites of HCV envelope proteins.** N-linked glycans are indicated by an N. Amino acid positions within the polyprotein correlate with HCV strain H. Glycans involved in HCV entry are symbolized by black squares. Mutated glycosylation sites, which alter folding of protein, are displayed with a grey circle. E1 and E2 ectodomains constructed for this work end with AA 326 and 661, respectively (indicated by arrows). Adapted from Lavie et al., 2007 101010.

For protein expression Schneider 2 (S2) cells, derived from *Drosophila melanogaster* were used. These cells grow as a loose, semi-adherent monolayer in tissue culture flasks, as suspension in spinners and shake flasks, at room temperature. They do not require CO_2. But the most important feature of these cells is their glycosylation pattern, which generally consists of high-mannose glycans,

as it is the case for E2[6]. Therefore, the Drosophila Expression System from Invitrogen was used to express and secrete HCV glycoprotein ectodomains into the supernatant of cultured S2 cells. To increase yield and quality of the proteins, the expression vector, encodes for a BiP secretion signal upstream the HCV sequences. Expression is induced via copper sulfate due to a metallothionein (MT) promotor. A six-fold C-terminal His-tag replaces the transmembrane domain in both glycoproteins.

4.4.2 Transfection of S2 Cells with Calcium Phosphate

Cells (cultured in 10 % FCS) were seeded with a density of $1x10^6$ cells/ml, in a six-well plate (volume 3 ml). Cells were co-transfected approximately two days later, at a density of $2x10^6$ cells/ml with the plasmids $E2_{661}$xpMT/V5/BiP-His and pCoHygro selection vector. Transfection for stable expression of S2xE2e was performed according to the manufacturer's instructions with these amounts and volumes:

tube A	36 µl $CaCl_2$ x µl DNA (19 µg) x µl pCoHygro (1 µg) ad 300 µl H_2O_{dest}
tube B	300 µl 2X HBS

24h after transfection, cells were washed twice with medium; induction of expression was induced with 5 mM SO_4Cu. After five days of incubation, cells were settled to cell culture flasks (T25), split in a 1:1 ratio and selection started with 300 µg/ml hygromycin. Additionally, reduction of FCS to zero was started. Analysis of expression was performed via Western-blot with α-E2 and/or α-His antibody five days later.

4.4.3 Competition Assay

To analyze correct functionality of the expressed E2 ectodomain, a competition assay was performed. For this, Huh7.5 cells were seeded at least four hours pre incubation with different supernatants, in different dilutions. After one-hour incubation, cells were additionally infected with a reporter virus. Eight hours later, medium was exchanged again. Cells were analyzed three days post infection via luciferase assay. If the E2 ectodomain is functional, it binds to HCV receptors located on the host cell, and therefore hampers HCV entry.

4.4.4 Purification of Supernatant from E2 Expressing Cells via HPLC

Preparative purification of the histidine tagged E2 ectodomain was performed via IMAC (immobilized metal ion affinity chromatography). For this, a HisTrap HP column with Ni Sepharose™ (1ml; GE Healthcare) in combination with the liquid chromatography system ÄKTA™ (ÄKTA explorer, GE Healthcare) was used.

Supernatants were harvested nearly every fifth day and centrifuged for 20 minutes at 4,000 rpm, to remove cells and debris. Before purification with HisTrap™ FF columns, samples were mixed 1:1 with binding buffer to prevent binding of host proteins with exposed histidines. At this step, precipitates accumulated, which could be dissolved by lowering the pH a bit with HCl, until sample is bright again. To prevent column clogging, samples had to be filtered through a 0.2 µm filter and buffers have to be degased additionally.

Binding buffer: 20 mM sodium phosphate; 0.5 M NaCl; 40 mM imidazole; pH = 7.4

Elution buffer: 20 mM sodium phosphate; 0.5 M NaCl; 500 mM imidazole; pH = 7.4

Next steps were performed, following manufacturer's instructions regarding the column and the ÄKTA system. The column has to be washed with five column-volumes (CVs) distilled water, followed by equilibration with five CVs binding buffer with a flow rate of 1 ml/min. Consequently 50 ml sample was applied via the injection valve and a syringe into a superloop. Via this superloop the sample was loaded onto the column with 1 ml/min. After loading, column had to be washed with binding buffer until absorbance reached a steady state baseline (approximately after 20 CVs). Elution could now be started with an increasing gradient of elution buffer until gradient reached 100 %, over 20 CVs with a flow rate of 1 ml/min, to separate proteins with similar binding strengths. Fraction volume was 2 ml. Fractions with the highest absorbance could be analyzed further onto SDS-PAGE followed by Coomassie staining and/or Western-blot.

The column could be washed after usage with elution buffer, distilled water and 20 % EtOH (about 20 CVs each, 2 ml/min) and stored at 4°C.

Since stripping of Ni-Ions from the column appeared within some purifications, presumably due to additional components in the supernatant, also HisTrap™ Excel columns were used where stripping is no longer possible and no further clarification than centrifugation is needed. For optimal binding no imidazole is recommended in sample and equilibration buffer.

Equilibration buffer: 20 mM sodium phosphate; 0.5 M NaCl; pH = 7.4

Wash buffer: 20 mM sodium phosphate; 0.5 M NaCl; 0 to 30 mM imidazole; pH = 7.4

Elution bufffer: 20 mM sodium phosphate; 0.5 M NaCl; 500 mM imidazole; pH = 7.4

Next steps were performed, following manufacturer's instructions regarding the column and the ÄKTA system. The column has to be washed with five CVs distilled water, followed by equilibration with five CVs equilibration buffer with a flow rate of 1 ml/min. Consequently the sample was applied via the injection valve and a syringe into a superloop (50 ml). Via this superloop the sample was loaded onto the column with 1 ml/min. After loading, column had to be washed with wash buffer until absorbance reached a steady state baseline (approximately after 20 CVs). Elution could now be started with an increasing gradient of elution buffer until gradient reached 100 %, over 20 CVs with a flow rate of 1 ml/min. Fraction volume was 2 ml. Fractions with the highest absorbance could be analyzed further onto SDS-PAGE.

The column could be washed after usage with elution buffer, distilled water and 20 % EtOH (20 CVs, 2 ml/min) and stored at 4°C.

4.4.5 Gel Filtration

For gel filtration the Superdex 200 column (GE Healthcare) was used. This column is optimal for the separation of globular proteins between 10 and 600 kDa. Sample had to be concentrated in the buffer: 50 mM phosphate; 150 mM NaCl; pH = 7.5 down to 500 µl. Next steps were performed, following manufacturer's instructions regarding the column and the ÄKTA system. The column was washed with at least 50 ml distilled water and equilibrated with at least 50 ml phosphate buffer with a flow rate of 0.5 ml/min. After sample injection to the 500 µl loop, sample was loaded onto the column with a flow rate of 0.5 ml/min. Elution took place with the phosphate buffer in 500 µl fractions.

4.4.6 Ion Exchange – Cation Exchange

Separation within this method is based on the proteins charge. Cation exchange retains cations, since the column matrix contains anions. The sample from Ni-NTA purification has to be sterile filtrated to prevent clogging of the column, and two buffers are needed:

I Phosphate buffer 50 mM, pH = 7

II Phosphate buffer 50 mM + 1 M NaCl, pH = 7

The column UNO^{TM} S-6 (Bio-Rad) was washed with 5 CVs distilled water and 5 CVs phosphate buffer I (1 ml/min). Ni-NTA purified (sterile filtered) sample was loaded onto the column (via 5 ml loop) and washed with 6 CVs phosphate buffer I. Elution took place with increasing gradient to 0.5 M NaCl via the phosphate buffer II over 60 CVs. Hereafter, the gradient had to be raised to 1.0 M NaCl over 4 CVs and hold for 4 CVs before re-equilibrating the column with 5 CVs phosphate buffer I. The flow rate was 2 ml/min.

5 Results

This thesis has focused on mapping protein-protein interactions (PPIs) of HCV and the structural analysis of HCV envelope glycoproteins. A major part was to conduct a medium-throughput screen via FACS based FRET / FCET. In addition, I established an expression system to elucidate the three-dimensional structure of HCV E1 and E2.

5.1 HCV Interactome

A variety of novel intra-protein interactions of HCV was discovered and the network of HCV proteins was defined for the first time in living liver cells. The gained network provides clues about functions of HCV proteins and can guide to new experimental approaches and present novel targets for antiviral therapy.

5.1.1 Single Transfections in HEK293T Cells

Initially, all HCV fusion protein expression plasmids had to be analyzed for correct localization, expression, and fluorescence intensity. Therefore, single transfections in HEK293T cells were performed, to indicate functionality. FACS analyses were carried out to calculate transfection efficiency and expression efficiency measured by the mean fluorescence intensity of the respective transfected fusions (Fig. 9 & Fig. 10). Subcellular localization was investigated by confocal microscopy to check if there is a specific protein distribution within the cell (Fig. 9 & Fig. 10). Finally Western-blot analysis was performed to confirm steady state protein expression levels and stability of the fusions (Fig. 11).

As indicated (Fig. 9 - Fig. 11) all fusions are expressed and show correct size as well as localization within the cell. Nevertheless, expression of CFP-NS4A was not detectable by Western-blot (data not shown), probably due to low transfection efficiency of 5 % (+/- 2.95, n = 5), the lowest among all CFP-fusions (compare with Fig. 10). However, CFP-NS4A could be detected by flow cytometry and confocal microscopy. Furthermore, the CFP-NS4A fusion protein localized to specific subcellular compartments, indicating that it is indeed functionally expressed (Fig. 10).

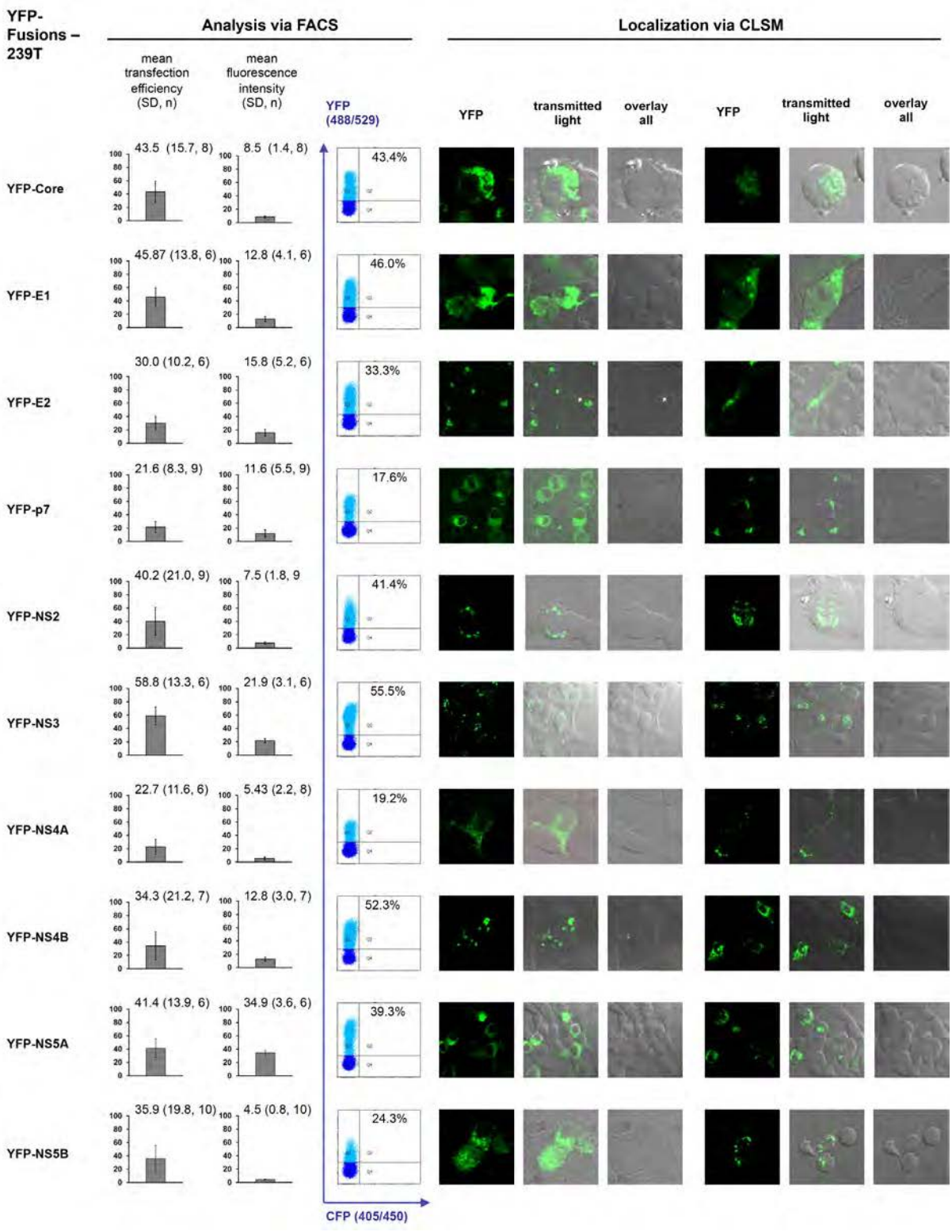

Fig. 9: **Expression & localization analysis of constructed YFP fusions via FACS and CLSM in HEK293T cells.** Detection of constructs was performed with BD FACSCantoII™ and Zeiss 510 Meta (Carl Zeiss, Jena). Mean transfection efficiency (mean, SD, n) and relative fluorescence intensity normalized to YFP adjusted to 100 % (mean, SD, n) and additionally one example for a single experiment with transfection efficiency (%) and YFP fluorescence intensity indicated via the y-axis is depicted. Localization of corresponding YFP-fusions was analyzed via confocal laser scanning microscopy (CLSM).

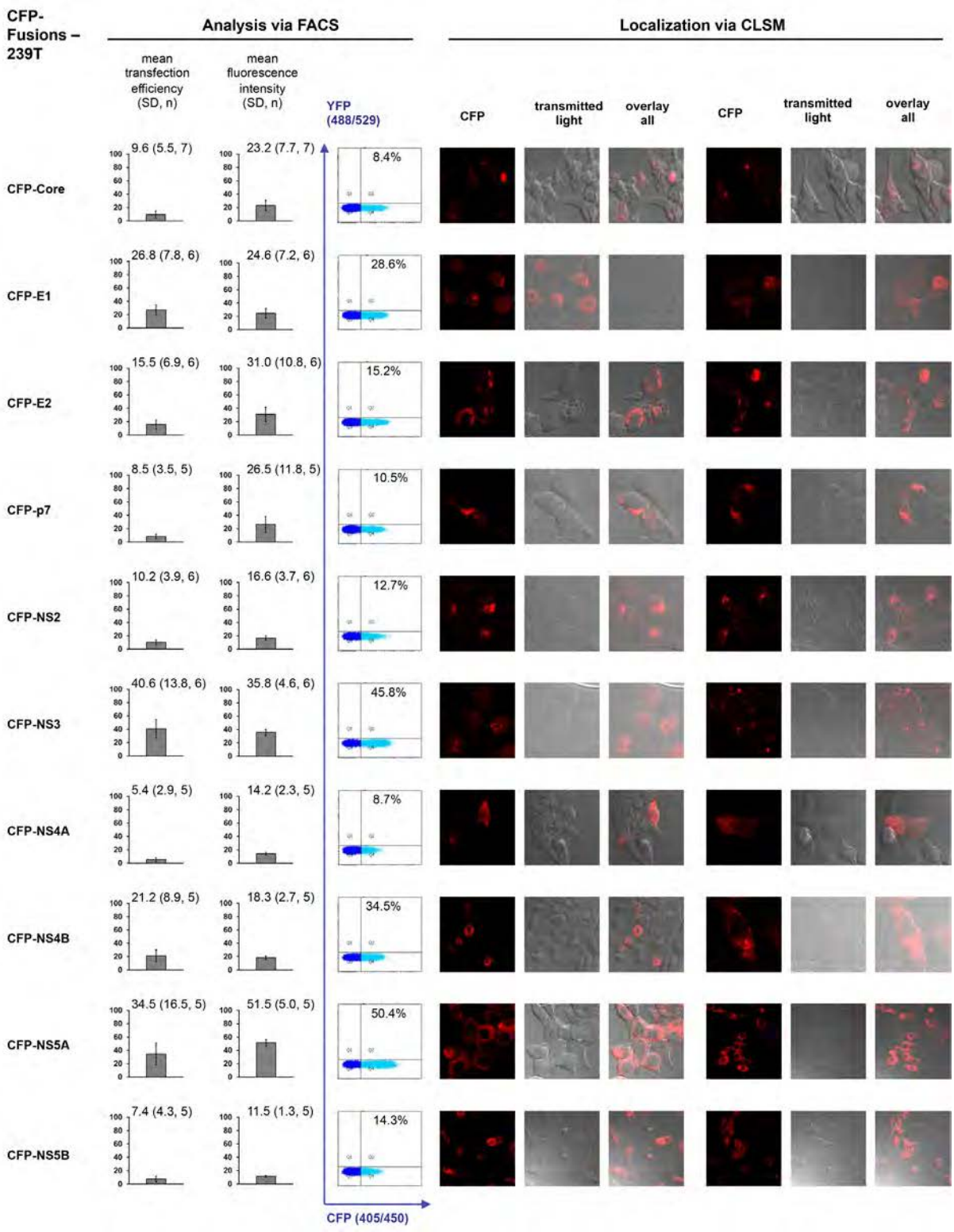

Fig. 10: **Expression & localization analysis of constructed CFP fusions via FACS and CLSM in HEK293T cells.** Detection of constructs was performed with BD FACSCantoII™ and Zeiss 510 Meta (Carl Zeiss, Jena). Mean transfection efficiency (mean, SD, n) and relative fluorescence intensity normalized to CFP adjusted to 100 % (mean, SD, n) and additionally one example for a single experiment with transfection efficiency (%) and CFP fluorescence intensity indicated via the x-axis is depicted. Localization of corresponding CFP-fusions was analyzed via confocal laser scanning microscopy (CLSM).

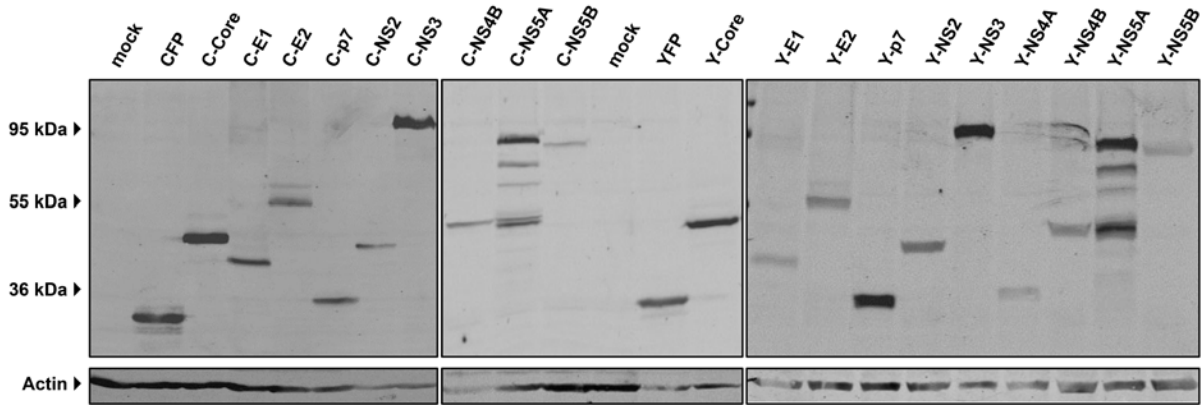

Fig. 11: **Expression analysis of constructed CFP and YFP fusions via Western-blot in HEK293T cells.** Detection of constructs was performed with anti-GFP antibody (1:10.000 diluted, BioVision), actin served as loading control. $1.2x10^5$ cells per lane were analyzed. Abbreviations: C, CFP; Y, YFP.

In the experiments shown in Fig. 9 to Fig. 11 all HCV proteins were fused with the chromophore on their N-terminus. Nevertheless, variants of Core, E1, and E2 with the chromophore fused to their C-terminus were also constructed and tested for transfection efficiency, and fluorescence intensity. The localization of the tag can have an impact on the functionality and the FRET signal. However, in contrast to the N-terminal tagged fusions, all tested C-terminal tagged HCV proteins were not functionally expressed and did neither show a proper subcellular localization or a pronounced transfection efficiency and fluorescence intensity (data not shown). Therefore, all HCV fusion proteins used in this study are N-terminal tagged with ECFP or EYFP.

5.1.2 Co-Transfections in HEK293T Cells

Next, all fusion proteins were tested for interaction by FACS based FRET in two combinations: YFP-protein A with CFP-protein B and CFP-protein A with YFP-protein B. Extensive FRET measurements and analysis of co-localization were performed after co-transfection of HEK293T cells. Obtained results were classified according to their FRET percentage: strong FRET signals, with FRET ≥ 25 % (I), FRET signals with medium range of 10 to 25 % (II), low signals from 2 to 10 % (III) and pairs which gave no FRET signal at all (IV; 0 – 2 %). Classification as no FRET signal was according to background signals of the negative control (CFP-Fusion transfected with YFP alone), which usually were in the range of 0 – 2 % (average value 0.49 %; +/- 0.93; n = 190). In the same manner, confocal pictures were classified, based on a complete co-localization of both transfected fusions (I), partial co-localization (II) and no co-localization at all (III). All data is summarized in Fig. 22.

A partial co-localization can be seen e.g. for CFP-NS2/YFP-E1 und CFP-NS3/YFP-p7 (Fig. 16 & Fig. 17). A complete overlap of the protein distribution pattern is seen amongst others for CFP-NS5A/YFP-

NS5A and CFP-NS5A/YFP-p7 (Fig. 20). An example for no co-localization represents the transfected combination CFP-Core/YFP-NS4B (Fig. 12).

Due to different expression levels and different fluorescence intensities, no specific adjustment of transfected DNA amounts was performed. However, the ratio of transfected CFP plasmid versus YFP plasmid was 1.5:1 to circumvent the overall lower fluorescence intensity of CFP in comparison to YFP.

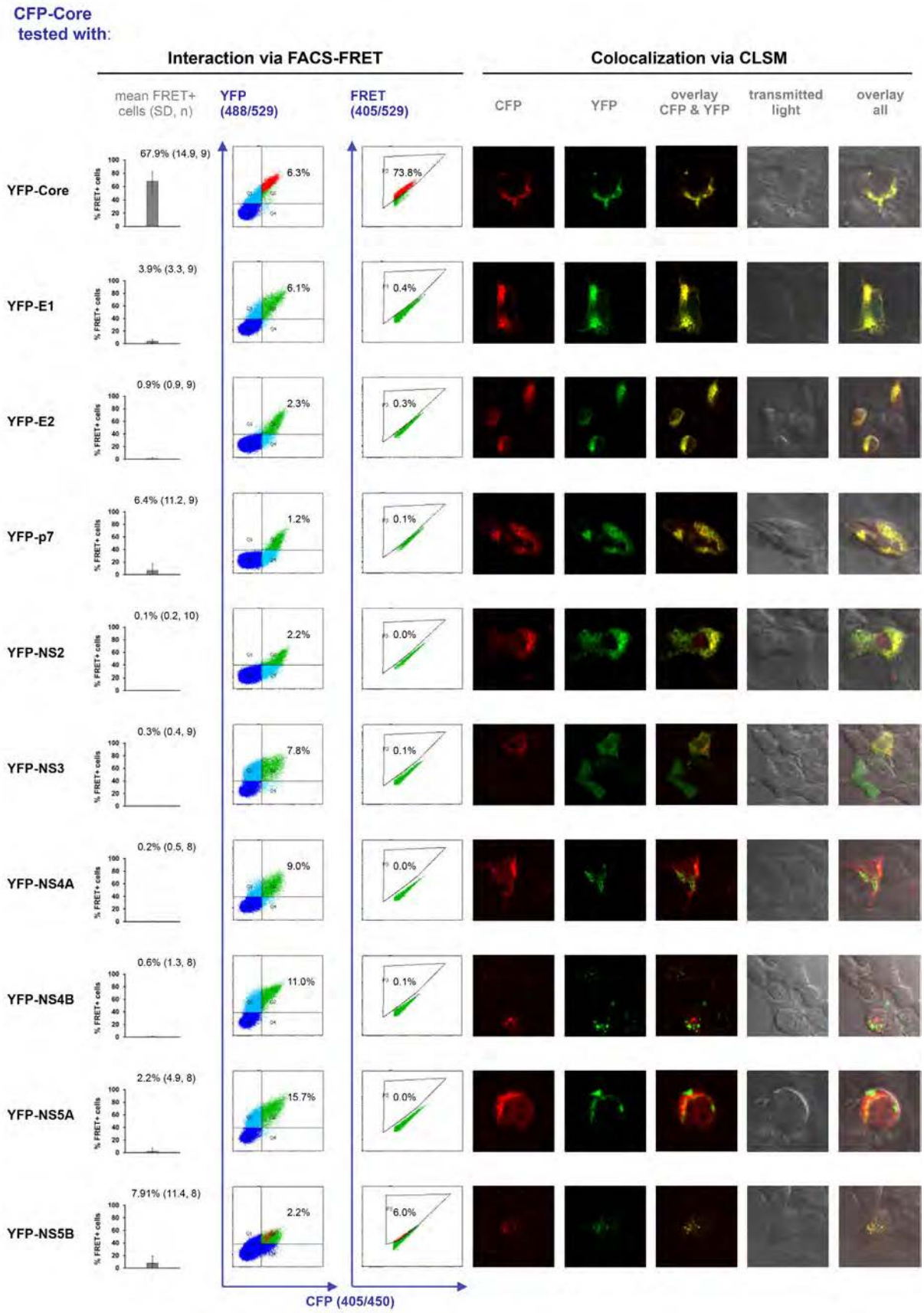

Fig. 12: **CFP-Core interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-Core. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

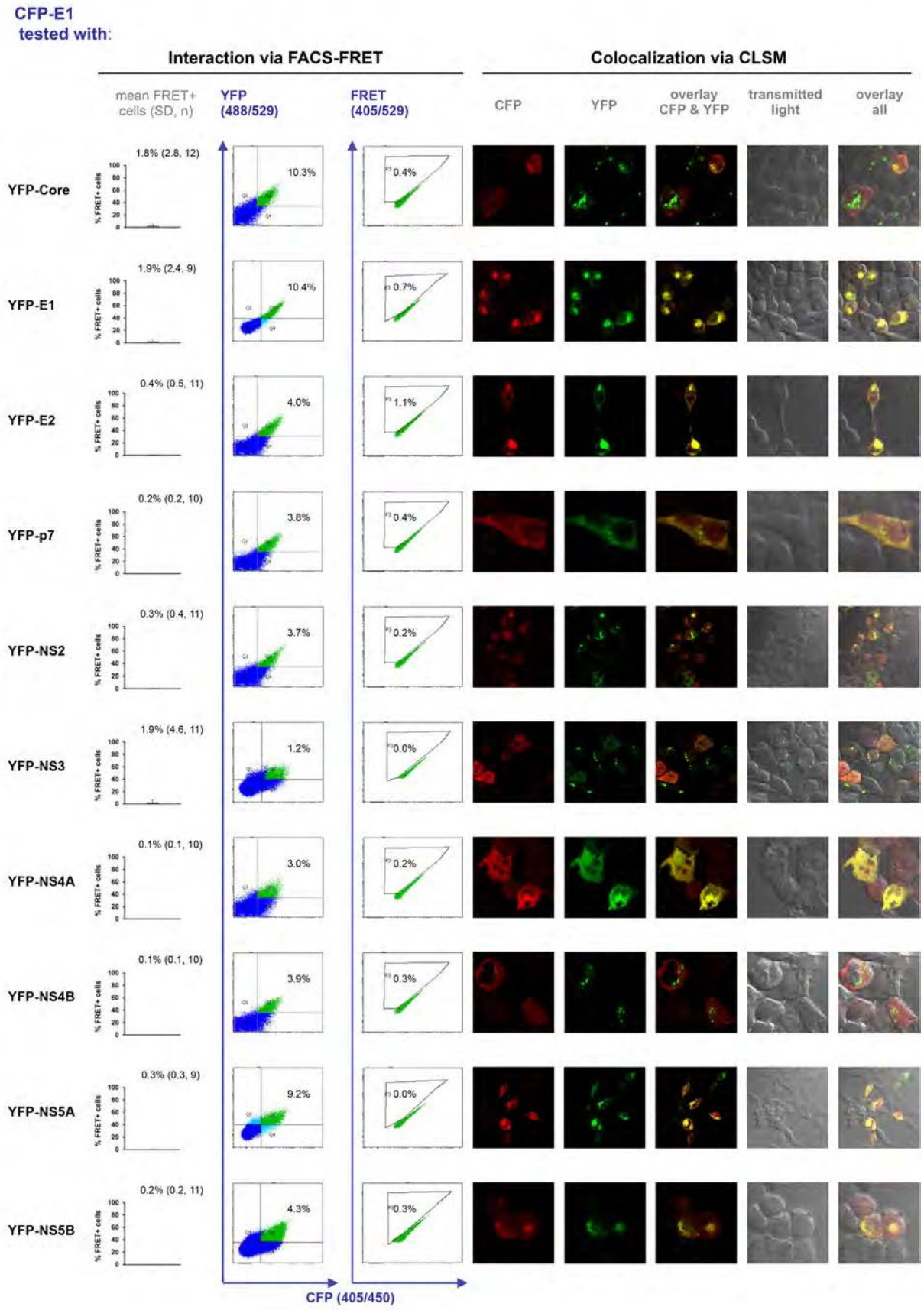

Fig. 13: **CFP-E1 interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-E1. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

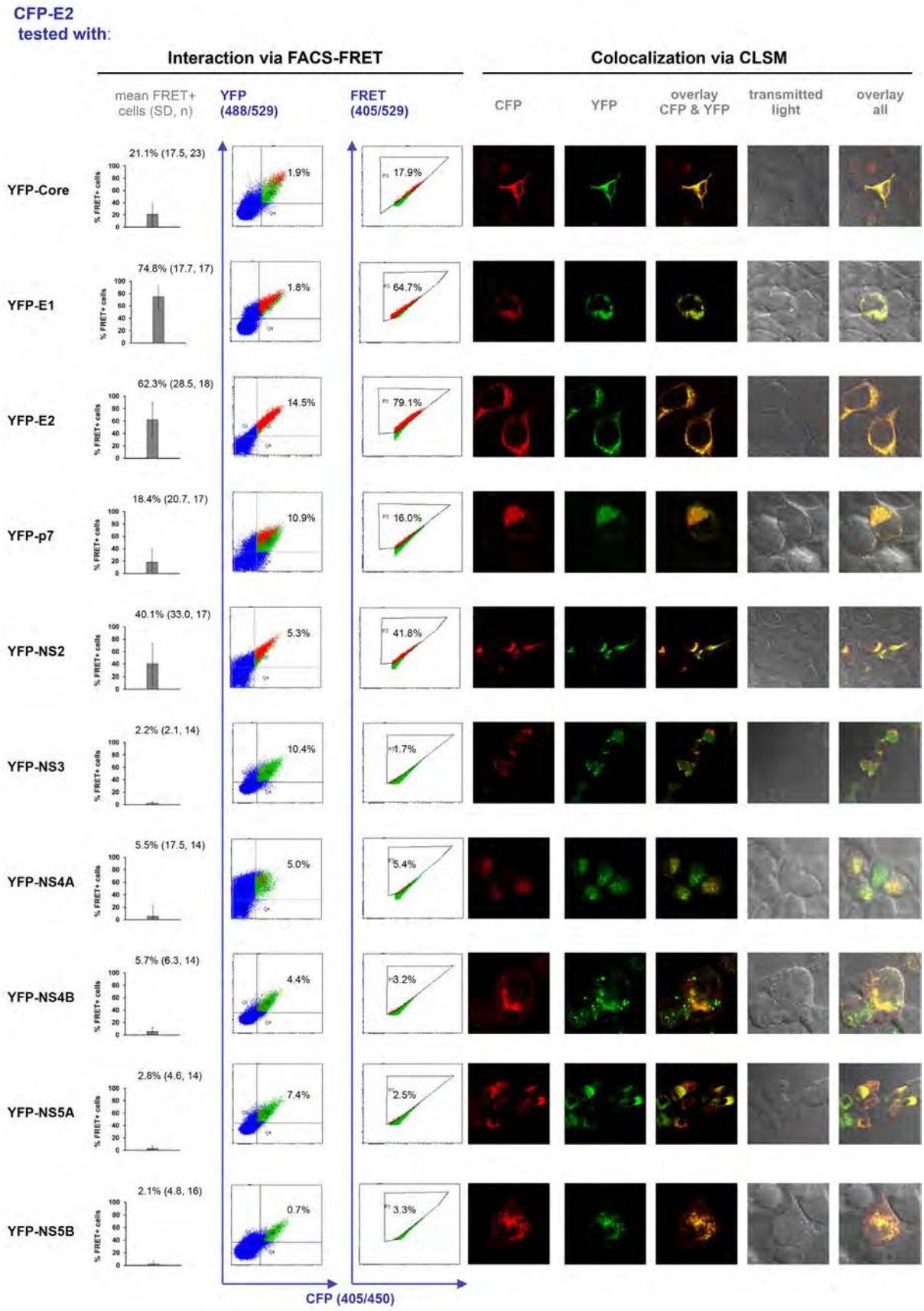

Fig. 14: **CFP-E2 interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-E2. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

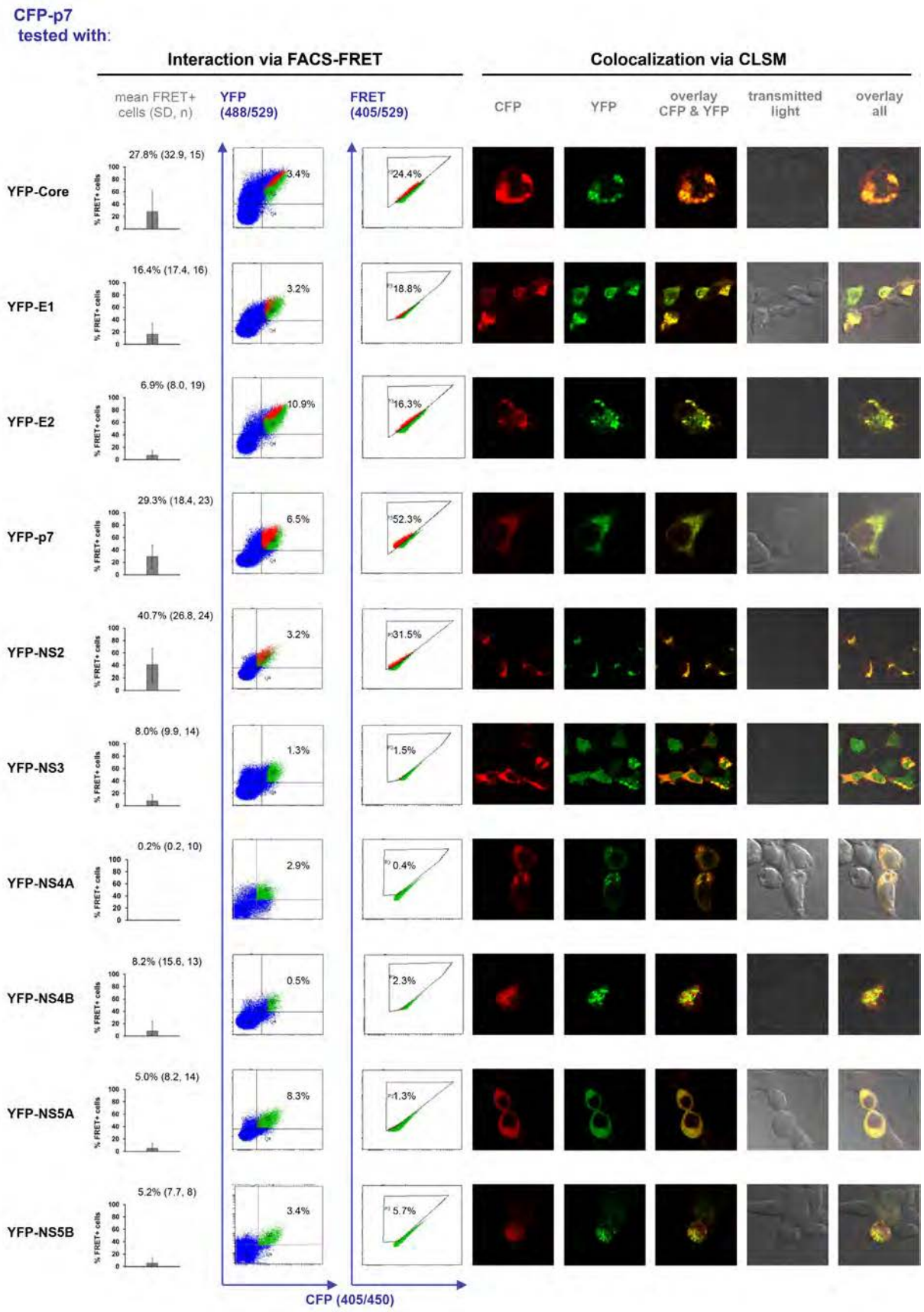

Fig. 15: **CFP-p7 interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-p7. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

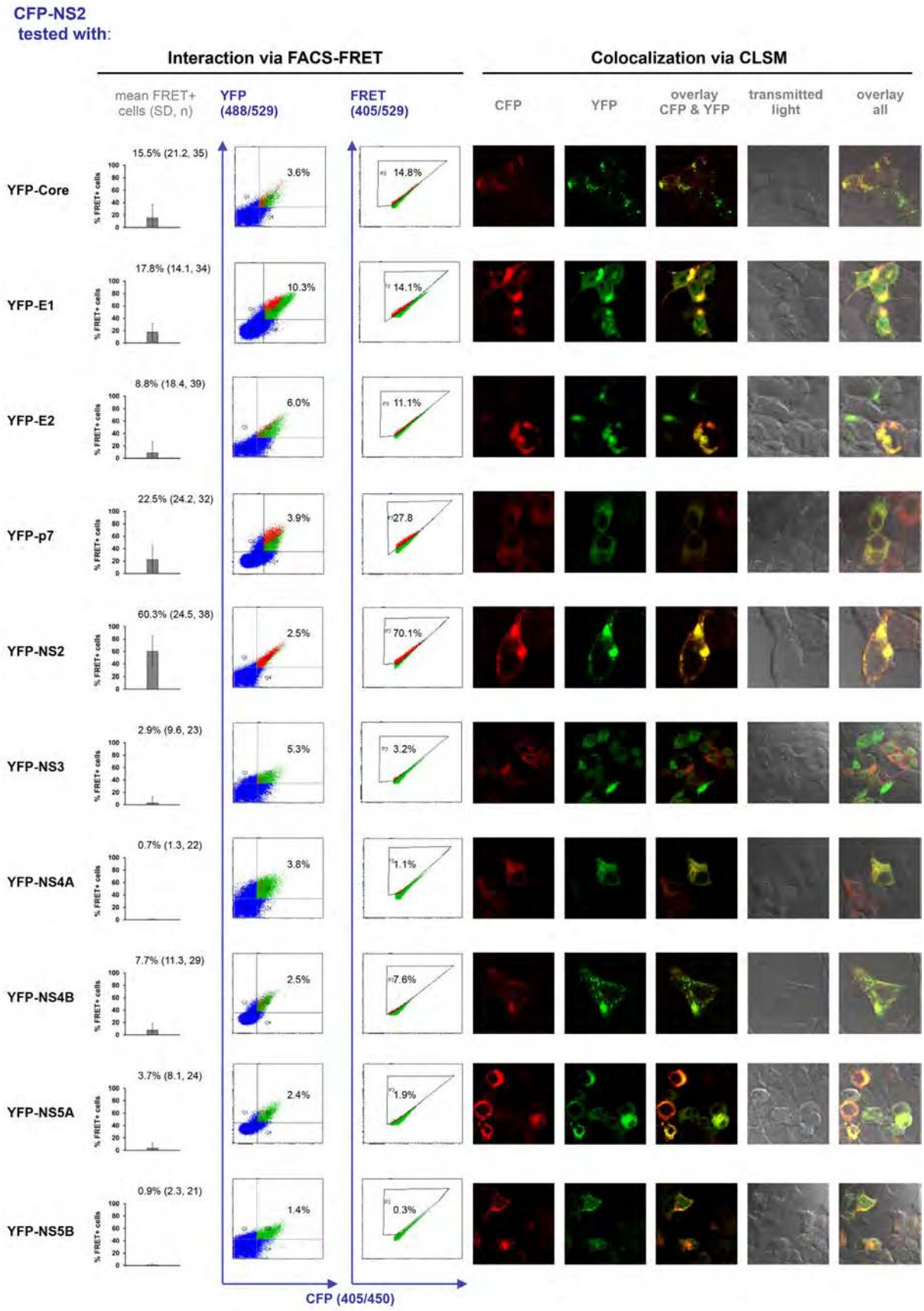

Fig. 16: **CFP-NS2 interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-NS2. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

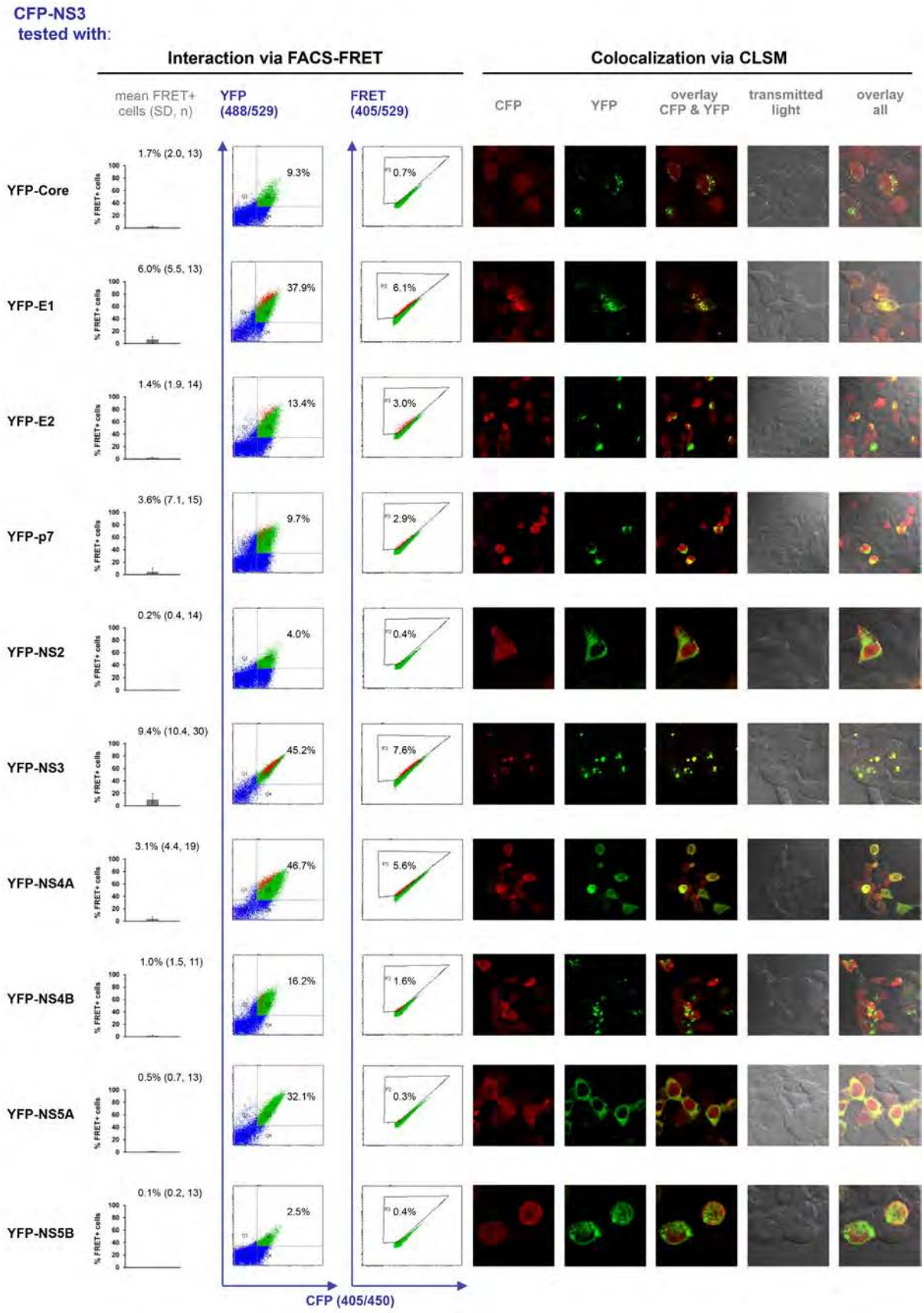

Fig. 17: **CFP-NS3 interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding C-NS3. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

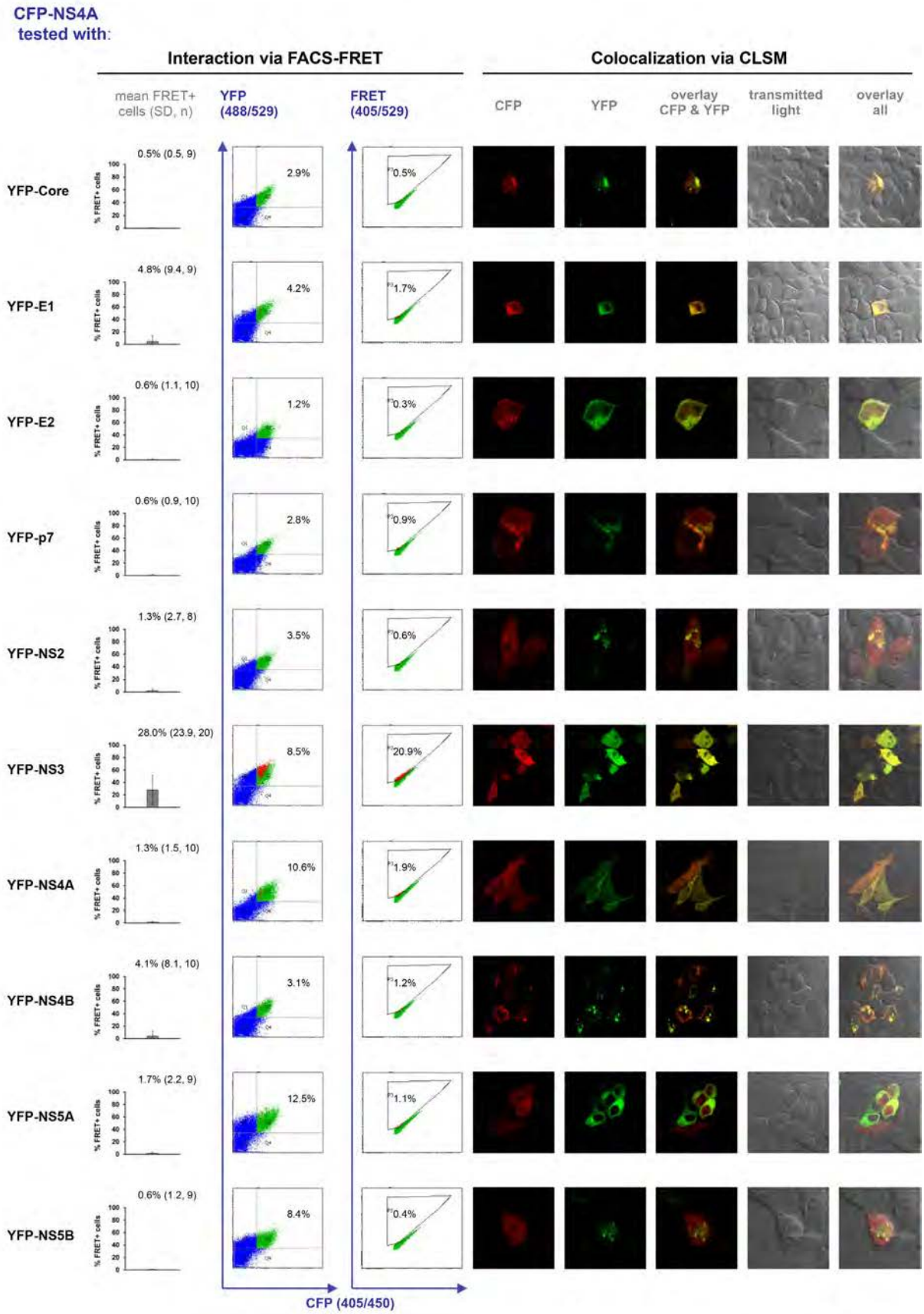

Fig. 18: **CFP-NS4A interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-NS4A. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

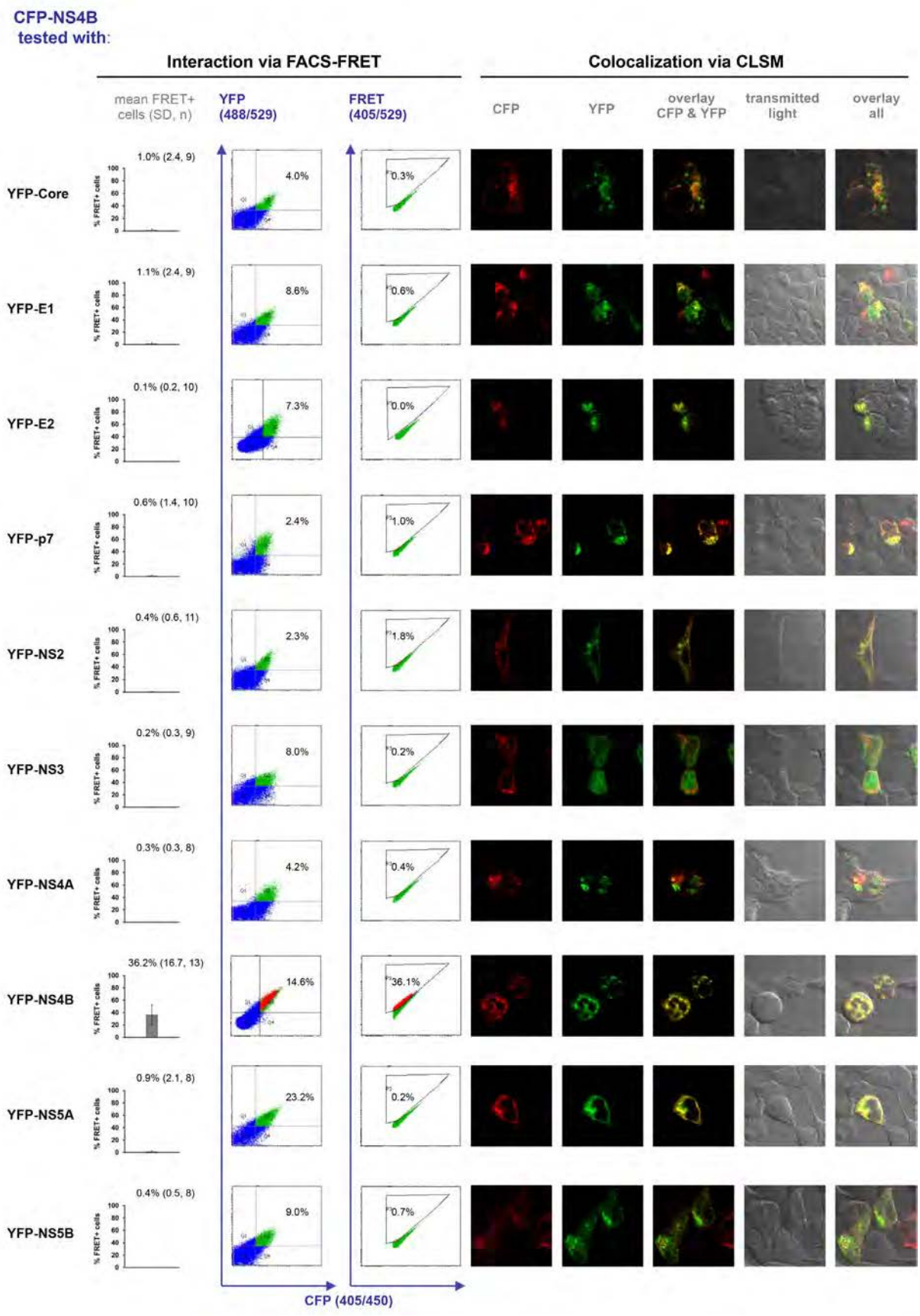

Fig. 19: **CFP-NS4B interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-NS4B. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

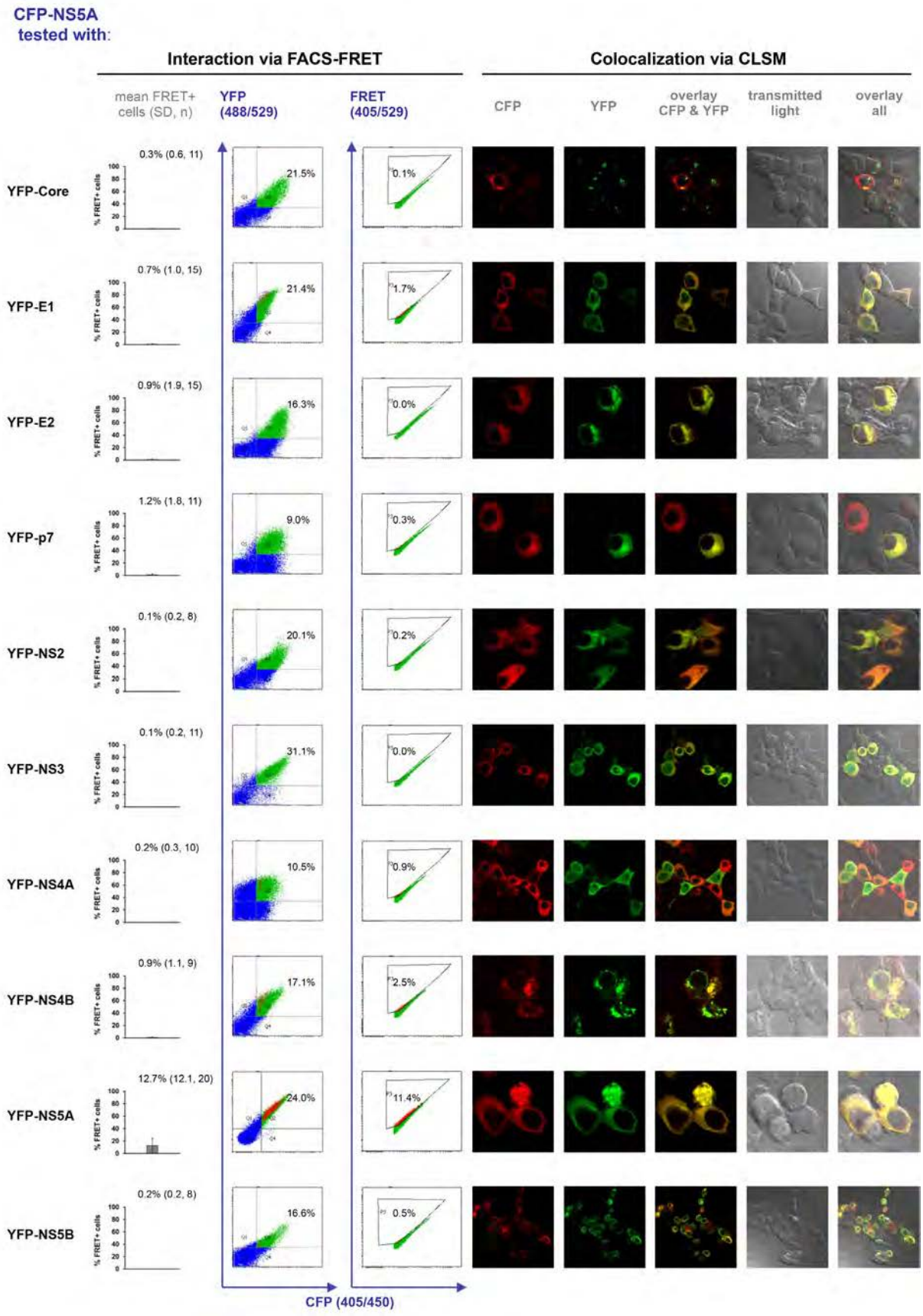

Fig. 20: **CFP-NS5A interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-NS5A. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

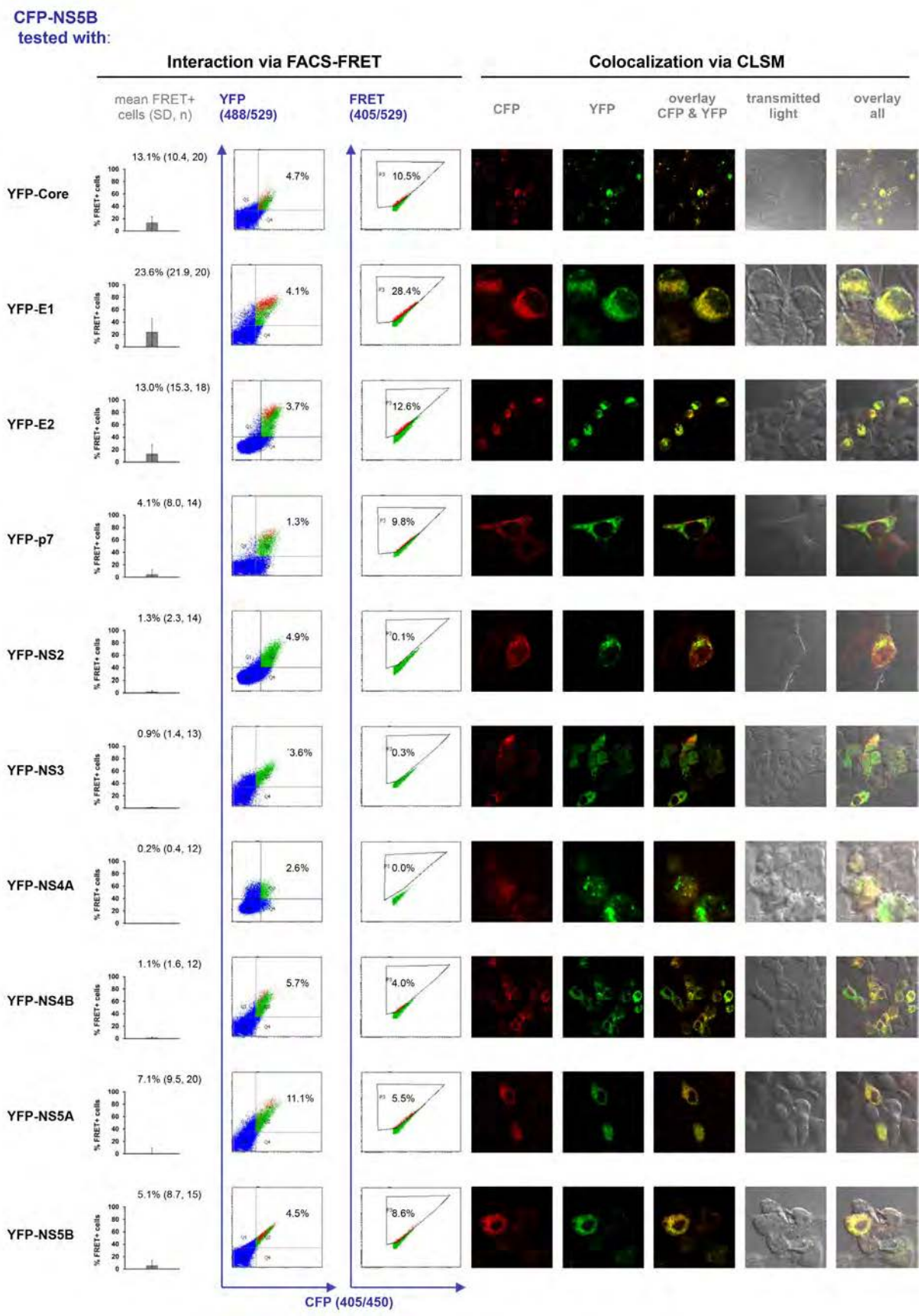

Fig. 21: **CFP-NS5B interplay with YFP fusions in HEK293T cells.** Indicated FRET (mean, SD, n) for all performed co-transfections regarding CFP-NS5B. Single examples show co-transfection efficiency (%, CFP against YFP) and FRET (%, CFP against FRET). Co-transfected cells are depicted by green and red dots. FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

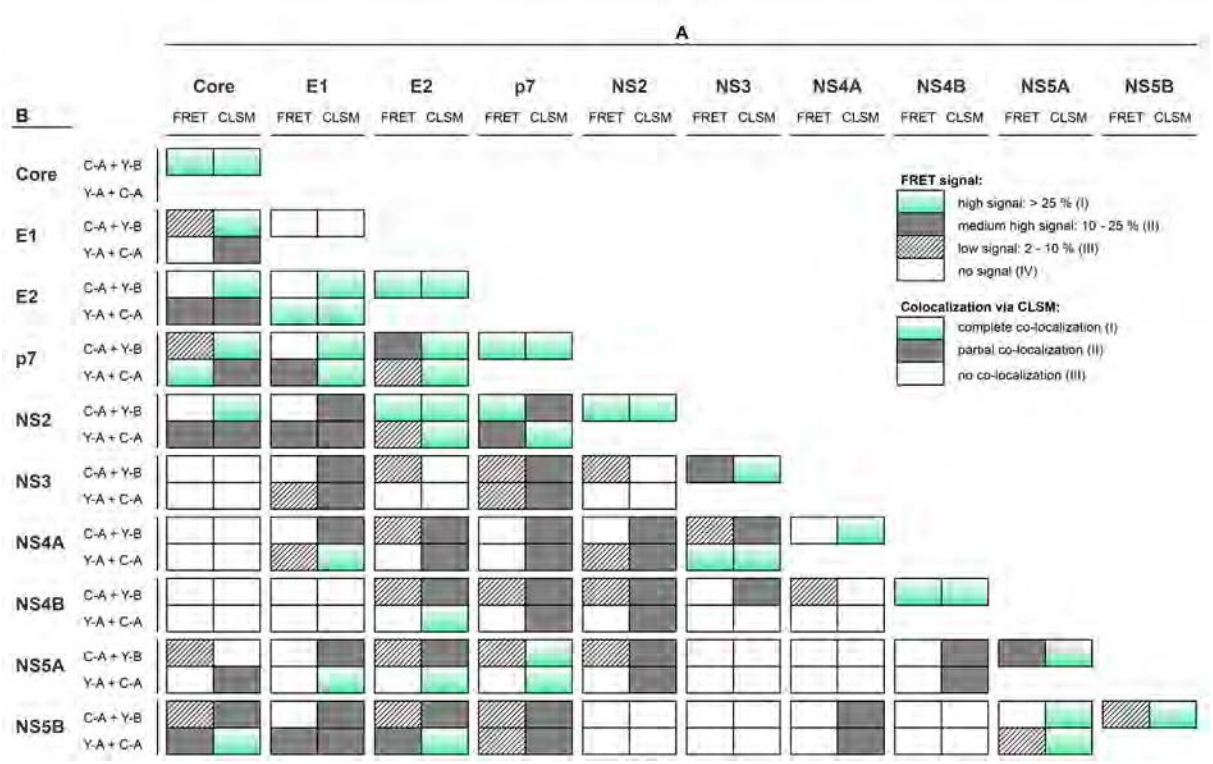

Fig. 22: **Overview of FRET and co-localization results in transfected HEK293T cells.** Indicated are four classes (I to IV) of different FRET signal ranges and three classes (I to III) of co-localization. Co-transfected HEK293T cells were analyzed for interaction via FACS based FRET and for co-localization via confocal laser scanning microscopy (CLSM). The figure shows the results of FRET signals for both combinations (YFP-protein A + CFP-protein B & CFP-protein A + YFP-protein B). The levels for FRET ranges are: high FRET: all signals higher than 25 % (I); medium FRET: signals from 10 to 25 % (II); low signals from 2 to 10 % (III) and background signal (IV; 0 – 2 %). Similar classifications were introduced for co-localization studies via CLSM: positive / overall co-localization (I), partial co-localization (II) and no co-localization (III).

To show interactions of all HCV proteins revealed in this thesis, results of performed co-transfections, regarding their expression, and their interplay in HEK293T cells are listed in detail (Fig. 12 – Fig. 21). Due to this, Fig. 22 was generated to give a review of gained results. As expected, FRET did only occur in cells in which the fluorescently labeled proteins showed at least partial co-localization. Vice versa, no FRET can be detected if no co-localization is seen (CFP-Core with YFP-NS4A, YFP-NS4B and YFP-NS5A). Additionally, FRET does not always occur in both tested combinations (YFP-protein A + CFP-protein B & CFP-protein A + YFP-protein B), maybe due to stoichiometric reasons. This can be observed especially for CFP-E1, which did not show any positive FRET signal in co-transfections with the YFP-fusions at all. Therefore, it is recommended to test indeed both combinations to circumvent the loss of information. The majority of HCV proteins is able to build homomers.

Since some HCV protein interactions might be dependent on a complex and only become evident upon co-expression of an additional viral protein, some triple-transfections were randomly tested. For this purpose, fluorescently labeled HCV proteins were transfected as before. However, an additional non-fluorescently tagged HCV protein was co-transfected to mimic the presence of accessory HCV proteins in infected cells (Fig. 23). No significant differences were detected compared to co-transfections with just two fluorescently tagged proteins.

A

	Construct 1	Construct 2	Construct 3	MF	SD	n
Core	CFP-Core	YFP		0,40	0,34	9
	CFP-Core	YFP-Core		69,61	14,99	10
	CFP-Core	YFP-Core	Core	62,60	39,60	2
	CFP-Core	YFP-Core	E2	68,95	29,20	2
	CFP-Core	YFP-Core	p7	56,68	19,70	4
	CFP-Core	YFP-Core	NS2	65,38	19,50	4
	CFP-Core	YFP-Core	NS3	68,00	36,20	2
E2	CFP-E2	YFP		0,65	1,08	22
	CFP-E2	YFP-E2		63,43	28,17	19
	CFP-E2	YFP-E2	Core	79,20	14,99	2
	CFP-E2	YFP-E2	E2	83,15	12,37	2
	CFP-E2	YFP-E2	p7	80,50	13,58	2
	CFP-E2	YFP-E2	NS2	76,45	16,19	2
	CFP-E2	YFP-E2	NS3	82,30	15,27	2
p7	CFP-p7	YFP		0,51	1,13	21
	CFP-p7	YFP-p7		36,00	22,36	28
	CFP-p7	YFP-p7	Core	56,70		1
	CFP-p7	YFP-p7	E2	57,10		1
	CFP-p7	YFP-p7	p7	46,80		1
	CFP-p7	YFP-p7	NS2	53,78	4,86	5
	CFP-p7	YFP-p7	NS3	63,10		1
NS2	CFP-NS2	YFP		0,88	1,43	42
	CFP-NS2	YFP-NS2		60,66	23,50	43
	CFP-NS2	YFP-NS2	Core	51,30		1
	CFP-NS2	YFP-NS2	E2	50,10		1
	CFP-NS2	YFP-NS2	p7	41,06	3,70	5
	CFP-NS2	YFP-NS2	NS2	36,20		1
	CFP-NS2	YFP-NS2	NS3	39,70		1
NS3	CFP-NS3	YFP		0,22	0,39	22
	CFP-NS3	YFP-NS3		10,56	12,11	31
	CFP-NS3	YFP-NS3	Core	16,10	10,89	2
	CFP-NS3	YFP-NS3	E2	25,15	12,66	2
	CFP-NS3	YFP-NS3	p7	15,65	6,44	2
	CFP-NS3	YFP-NS3	NS2	9,95	5,73	2
	CFP-NS3	YFP-NS3	NS3	33,45	13,93	2
NS4A	CFP-NS4A	YFP		1,68	2,08	15
	CFP-NS4A	YFP-NS4A		0,94	1,31	15
	CFP-NS4A	YFP-NS4A	Core	1,30		1
	CFP-NS4A	YFP-NS4A	E2	1,50		1
	CFP-NS4A	YFP-NS4A	p7	1,00		1
	CFP-NS4A	YFP-NS4A	NS2	1,10		1
	CFP-NS4A	YFP-NS4A	NS3	0,40	0,63	5
NS4B	CFP-NS4B	YFP		0,29	0,38	13
	CFP-NS4B	YFP-NS4B		39,41	20,07	14
	CFP-NS4B	YFP-NS4B	Core	75,40		1
	CFP-NS4B	YFP-NS4B	E2	76,00		1
	CFP-NS4B	YFP-NS4B	p7	74,50		1
	CFP-NS4B	YFP-NS4B	NS2	75,00		1
	CFP-NS4B	YFP-NS4B	NS3	78,90		1
NS5A	CFP-NS5A	YFP		0,23	0,40	19
	CFP-NS5A	YFP-NS5A		13,96	13,18	21
	CFP-NS5A	YFP-NS5A	Core	29,20		1
	CFP-NS5A	YFP-NS5A	E2	37,80		1
	CFP-NS5A	YFP-NS5A	p7	33,50		1
	CFP-NS5A	YFP-NS5A	NS2	21,40		1
	CFP-NS5A	YFP-NS5A	NS3	40,40		1
NS5B	CFP-NS5B	YFP		0,67	1,31	21
	CFP-NS5B	YFP-NS5B		5,92	9,05	16
	CFP-NS5B	YFP-NS5B	Core	13,10		1
	CFP-NS5B	YFP-NS5B	E2	17,50		1
	CFP-NS5B	YFP-NS5B	NS2	11,80		1

triple:
CFP-A + YFP-B + C

B

	Construct 1	Construct 2	Construct 3	MF	SD	n
NS3/NS4A	CFP-NS3	YFP		0,22	0,39	22
	CFP-NS3	YFP-NS4A		3,12	4,39	19
	CFP-NS3	YFP-NS4A	NS2	0,60		1
	CFP-NS3	YFP-NS4A	p7	0,00		1
	CFP-NS3	YFP-NS4A	NS3	0,10		1
	CFP-NS3	YFP-NS4A	Core	0,20		1
	CFP-NS3	YFP-NS4A	E2	0,10		1
	CFP-NS4A	YFP		1,68	2,08	15
	CFP-NS4A	YFP-NS3		30,05	25,15	21
	CFP-NS4A	YFP-NS3	NS2	77,90		1
	CFP-NS4A	YFP-NS3	p7	78,10		1
	CFP-NS4A	YFP-NS3	NS3	71,60		1
	CFP-NS4A	YFP-NS3	Core	77,70		1
	CFP-NS4A	YFP-NS3	E2	69,90		1
	CFP-NS4A	YFP		1,68	2,08	15
	CFP-NS4A	YFP-Core		1,12	2,10	11
	CFP-NS4A	YFP-Core	NS3	2,85	3,61	2
	CFP-NS4A	YFP-E1		0,55	0,35	2
	CFP-NS4A	YFP-E1	NS3	0,35	0,35	2
	CFP-NS4A	YFP-E2		0,10	0,00	2
	CFP-NS4A	YFP-E2	NS3	0,10	0,14	2
	CFP-NS4A	YFP-p7		1,10	0,57	2
	CFP-NS4A	YFP-p7	NS3	0,45	0,50	2
	CFP-NS4A	YFP-NS2		0,95	1,34	2
	CFP-NS4A	YFP-NS2	NS3	0,30	0,42	2
	CFP-NS4A	YFP-NS4B		18,53	12,58	3
	CFP-NS4A	YFP-NS4B	NS3	23,70	21,90	3
	CFP-NS4A	YFP-NS5A		5,00	2,12	2
	CFP-NS4A	YFP-NS5A	NS3	5,45	1,06	2
	CFP-NS4A	YFP-NS5B		0,00	0,00	2
	CFP-NS4A	YFP-NS5B	NS3	0,00	0,00	2
	CFP-Core	YFP		0,18	0,35	4
	CFP-Core	YFP-NS4A		0,20	0,44	10
	CFP-Core	YFP-NS4A	NS3	0,05	0,07	2
	CFP-E1	YFP		0,24	0,39	10
	CFP-E1	YFP-NS4A		0,06	0,09	12
	CFP-E1	YFP-NS4A	NS3	0,05	0,07	2
	CFP-E2	YFP		0,65	1,08	22
	CFP-E2	YFP-NS4A		4,81	16,40	16
	CFP-E2	YFP-NS4A	NS3	0,00	0,00	2
	CFP-p7	YFP		0,20	0,34	17
	CFP-p7	YFP-NS4A		0,13	0,20	12
	CFP-p7	YFP-NS4A	NS3	0,00	0,00	2
	CFP-NS2	YFP		0,88	1,50	38
	CFP-NS2	YFP-NS4A		0,65	1,30	24
	CFP-NS2	YFP-NS4A	NS3	0,00	0,00	2
	CFP-NS4A	YFP		1,68	2,08	15
	CFP-NS4A	YFP-NS4A		0,94	1,36	14
	CFP-NS4A	YFP-NS4A	NS3	0,13	0,15	4
	CFP-NS4B	YFP		0,29	0,38	13
	CFP-NS4B	YFP-NS4A		0,24	0,29	10
	CFP-NS4B	YFP-NS4A	NS3	0,00	0,00	2
	CFP-NS5A	YFP		0,23	0,40	19
	CFP-NS5A	YFP-NS4A		0,13	0,26	12
	CFP-NS5A	YFP-NS4A	NS3	0,00	0,00	2
	CFP-NS5B	YFP		0,67	1,31	21
	CFP-NS5B	YFP-NS4A		0,16	0,36	14
	CFP-NS5B	YFP-NS4A	NS3	0,00	0,00	2

Fig. 23: **Randomized testing of triple transfections performed in HEK293T cells.** Depicted are on the left side triple-transfections regarding the HCV homomers and on the right side regarding the NS3/NS4A complex. Two fluorescently labeled constructs were co-transfected plus an additional non-tagged third construct (triple) to detect eventually more interactions due to complex formations and to mimic more the natural environment of HCV proteins in their host cell.

5.1.3 Co-Transfections in Huh7.5 Cells

HEK293T cells are kidney derived and were used for the FCET approach since they are an established and easy to transfect mammalian cell system that allows overexpression of proteins. Nevertheless, Hepatitis C virus preferentially targets liver cells. Thus, interactions of viral proteins could be different in the presence of liver cell specific factors. Therefore, transfection of Huh7.5 liver cells was established in a next step (see method section for details). All interactions found in HEK293T cells with FRET signals higher than 10 % (amongst some others below the threshold) were investigated in Huh7.5 cells in at least four independent transfections.

Background signals regarding negative controls for the Huh7.5 transfections showed an average of 0.29 % (+/- 0.97, n = 82). Results were classified in the same manner as described for the HEK293T cells before (see chapter 5.1.2).

Mean FRET values obtained in HEK293T and Huh7.5 cells were analyzed concerning the Pearson's correlation coefficient (Fig. 24). Although the absolute percentage of FRET-positive cells varied between both cell types, interactions found in HEK293T cells could generally be confirmed in the liver cell line (Pearson = 0.793; R^2 = 0.628; p < 0.0001; n = 45).

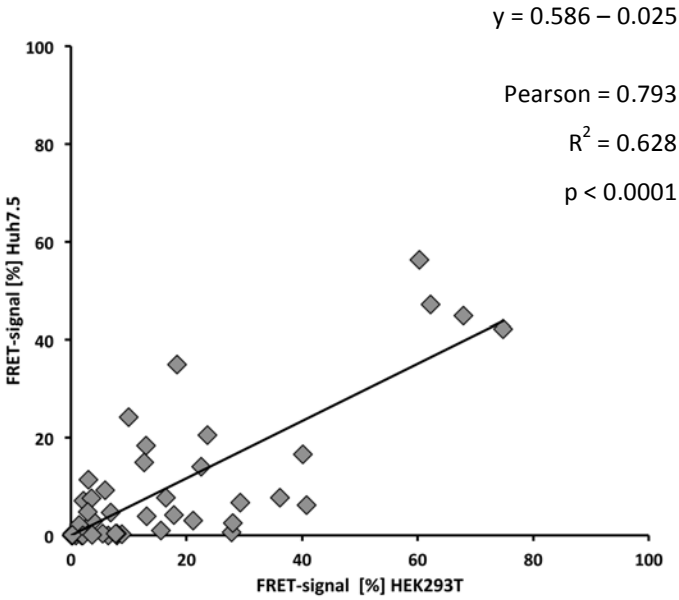

Fig. 24: **Correlation of FACS-FRET results generated in HEK293T and Huh7.5 liver cells** co-transfected with HCV fusion constructs: Pearson = 0.793, R^2 = 0.585; p < 0.0001.

For example found FRET signals for CFP-Core/YFP-E1 had nearly the same percentage in both tested cell lines (HEK293T: 3.91 %, +/- 3.33, n = 9 & Huh7.5: 2.87 %, +/- 3.53, n = 7). The same is true for the interplay of CFP-NS3/YFP-E1 (HEK293T: 5.96, +/- 5.51, n = 13 & Huh7.5: 9.30 %, +/- 7.54, n = 6) and CFP-NS5B/YFP-E1 (HEK293T: 23.62 %, +/- 21.94, n = 20 & Huh7.5: 20.55 %, +/- 5.03, n = 4). In contrast to that, dramatically differences can be seen for the CFP-p7/YFP-p7 interaction, where FRET signals in HEK293T cells are much higher (29.30 %, +/- 18.37, n = 23) than in the Huh7.5 cell line (6.69 %, +/- 6.69, +/- 4.66 n = 11). For the NS3/NS4A complex, differences are seen interestingly in both tested combinations. The combination CFP-NS3/YFP-NS4A shows in HEK293T cells very low signals (2.99 %, +/- 4.31, n= 20) in contrast to the Huh7.5 cell line (11.44, +/- 18.44, n = 8), but vice versa CFP-

NS4A/YFP-NS3 leads in HEK293T cells to very strong signals (27.97, +/- 23.88, n = 20) compared to the liver cell line (2.46, +/- 2.65, n = 14).

In general, expression of fusion proteins was more efficient in HEK293T cells. Thus, lower percentage of FRET signal in Huh7.5 cells could be due to less efficient expression of the fluorescent proteins.

5.1.4 Statistical Analysis of FRET Results in Both Cell Lines

Statistical analyses were performed using the Graph Pad Prism software (v. 5 for Mac). For all calculations the two-tailed unpaired Student's t-test was used to statistically challenge observed differences between FRET and background signals (respective CFP-fusion protein co-transfected with YFP-only). The 10 % threshold is user-defined, and was introduced as an additional stringency threshold for interactions. Thus, some FRET signals below the threshold are significant higher than the background, but are considered here as negative (see Fig. 25, green numbers).

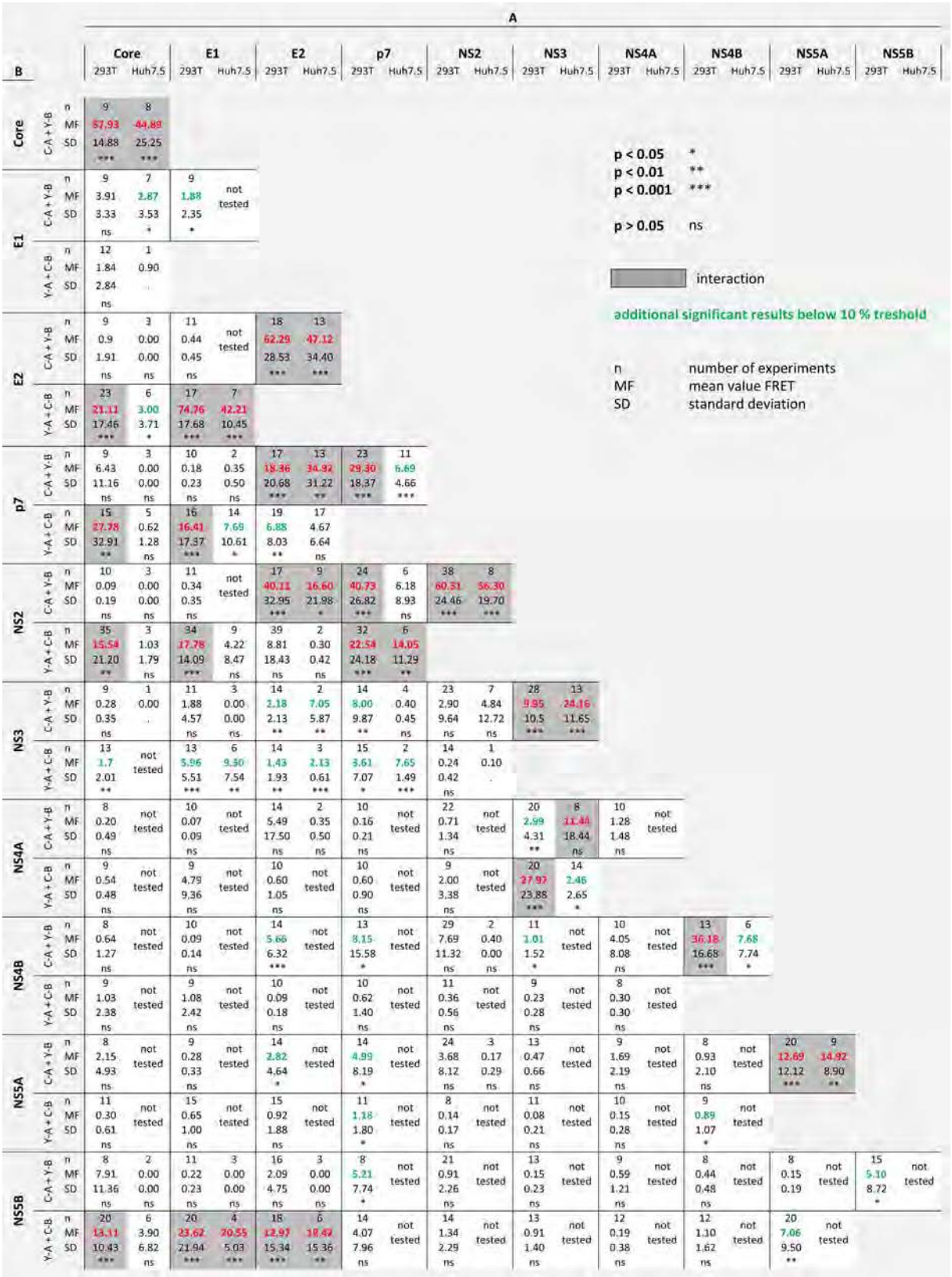

Fig. 25: **Statistical analyses of FRET signals in HEK293T and Huh7.5 cells.** Number of independent experiments (n), mean FRET values (MF), and standard deviations (SD) are indicated. Analyses were performed for both combinations (YFP-protein A + CFP-protein B & CFP-protein A + YFP-protein B), using the Graph Pad Prism software version 5 (Mac).
For calculation, FRET signals of n experiments were compared to the negative control FRET signal (CFP-fusion with YFP-only) with the two-tailed unpaired Student's t-test. Significance levels are indicated as follows: $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***).
Grey boxes depict FRET signals ≥ 10 % (numbers in red). This stringency criterion was empirically introduced to define interaction of co-transfected proteins. The analyses revealed that some mean FRET values below 10 % reached statistical significance. These values are shown in green.

Shown in Fig. 26 is a consolidated table of results found within the HEK293T and Huh7.5 screen, plus already postulated interactions from other groups. An x indicates newfound interactions. 12 out of 20 found interactions could be detected in both cell lines (Core/Core, E1/E2, E1/NS5B, E2/E2, E2/p7, E2/NS2, E2/NS5B, p7/NS2, NS2/NS2, NS3/NS3, NS3/NS4A, NS5A/NS5A) and eight exclusively in HEK293T cells (Core/E2, Core/p7, Core/NS2, Core/NS5B, E1/p7, E1/NS2, p7/p7, NS4B/NS4B). Seven found interactions are reported here as new, mainly regarding the structural HCV proteins (see also Fig. 26).

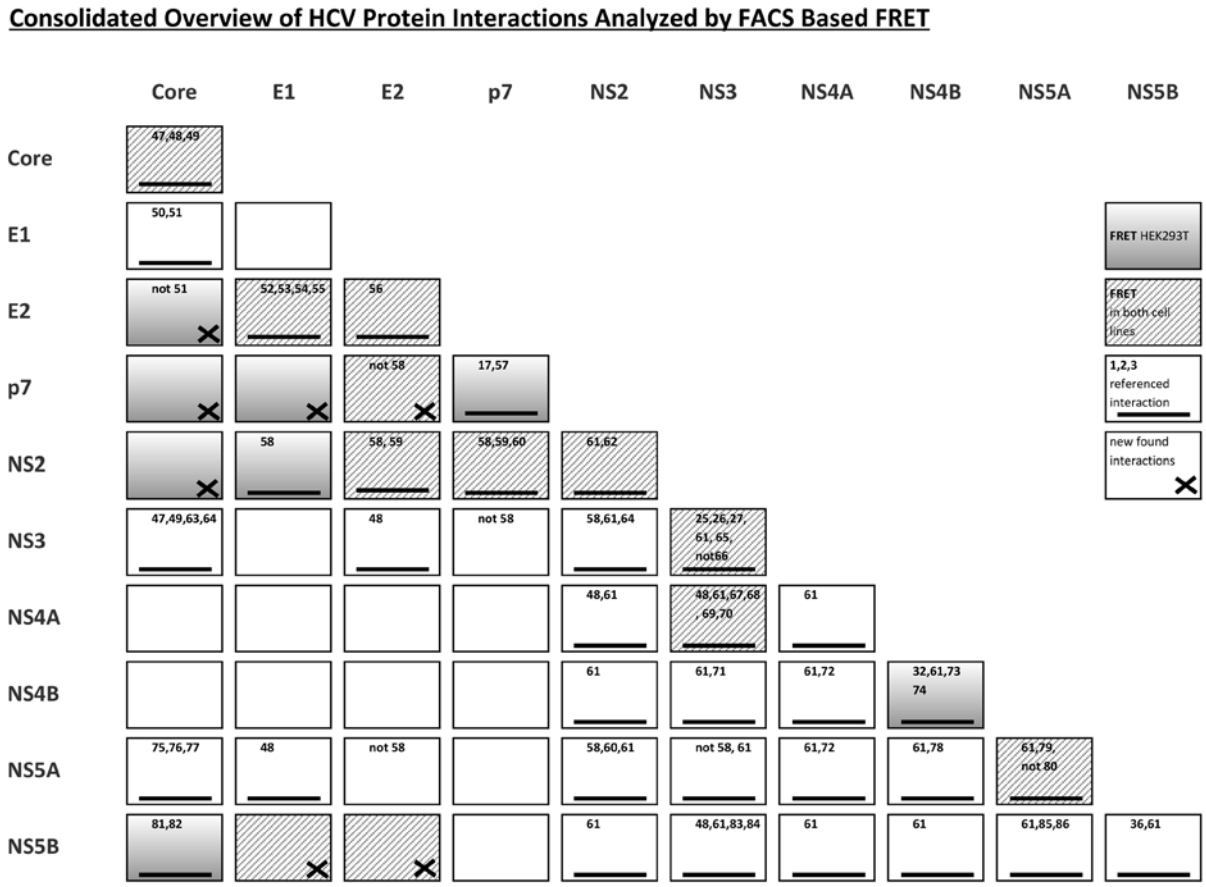

Fig. 26: **Overview of HCV protein interactions measured via FACS-FRET in both tested cell lines.** Statistical significant interactions with FRET values ≥ 10 % are presented. Interactions measured in HEK293T are highlighted by filled grey boxes and in Huh7.5 by striped grey boxes. Furthermore, already described interactions by literature are indicated (black undersCored bar, numbers for corresponding references are written in the box). Novel interaction reported within this thesis are indicated by an **x**.

5.1.5 Detailed Analyzes of Specific Interactions Revealed by FACS-FRET

5.1.5.1 Homomerization Amongst HCV Proteins

HCV proteins known to dimerize, often show very high FRET signals around 60 % in HEK293T cells. For example Core (67.93 %; +/- 14.88; n = 9), E2 (62.29 %; +/- 28.53; n = 18; see Fig. 28) as well as

NS2 (60.31 %; +/- 24.46; n = 38). Signals in liver cells are lower but still between 40 and 50 %. NS4B, which induces the membraneous web upon dimerization, has a lower but still high FRET signal of 36.18 % (+/- 16.68, n = 13) in HEK293T cell. In Huh7.5 cells the signal decreases under the 10 % threshold (7.68 %, +/- 7.74, n = 6). This could also be due to less efficient protein expression in the liver cell line.

Regarding NS3 homo-oligomerization (most likely dimerization[65]), FRET signal in the liver cells (24.16 %, +/- 11.65, n = 13) is more than twice as high as in kidney cells (9.95 %, +/- 10.5, n = 28). The already known interplay of serine protease NS3 with its cofactor NS4A could be determined in HEK293T (27.97 %, +/- 23.88, n = 20) and Huh7.5 cells (11.44 %; +/- 18.44; n = 8).

The phosphoprotein NS5A, known to affect inter alia the cellular stress response, shows self-reacting activity with a FRET signal of 12.69 % (+/- 12.12, n = 20) in HEK293T and 14.92 % (+/- 8.9, n = 9) in Huh7.5 cells.
It is noteworthy that FRET signals for p7 interplay, most likely the association of six[17] or even seven[57] molecules, in HEK293T cells show lower FRET intensity than for the dimers (29.30 %; +/- 18.37; n = 23). Energy transfer seems to be not efficient enough in Huh7.5 cells (6.69 %, +/- 4.66, n = 11).

5.1.5.2 Interplay of the HCV Glycoproteins

HCV glycoproteins might share some functionality with other envelope proteins of viruses belonging to the *Flaviviridae* family. In most cases the first envelope protein acts as a chaperone for the second, while the second acts as membrane fusion protein. A monomeric state is assumed for the chaperone, which I confirmed by absence of FRET with signals of 1.88 % (+/- 2.35, n = 9), far below the threshold of 10 %. However, the membrane fusion-protein E2 exists most likely as a dimer, which forms a trimeric state while it fuses with the host cell membrane. The percentage of FRET-positive cells for E2 multimerization (62.29 %; +/- 28.53; n = 18; see Fig. 28) supports such a mechanism. Also E1/E2 interaction with a FRET signal of 74.76 % (+/- 17.68, n = 17) supports the chaperone interplay of E1 with E2 in HEK293T and in Huh7.5 cells (42.21 %, +/- 10.45, n = 7). See details in Fig. 27.

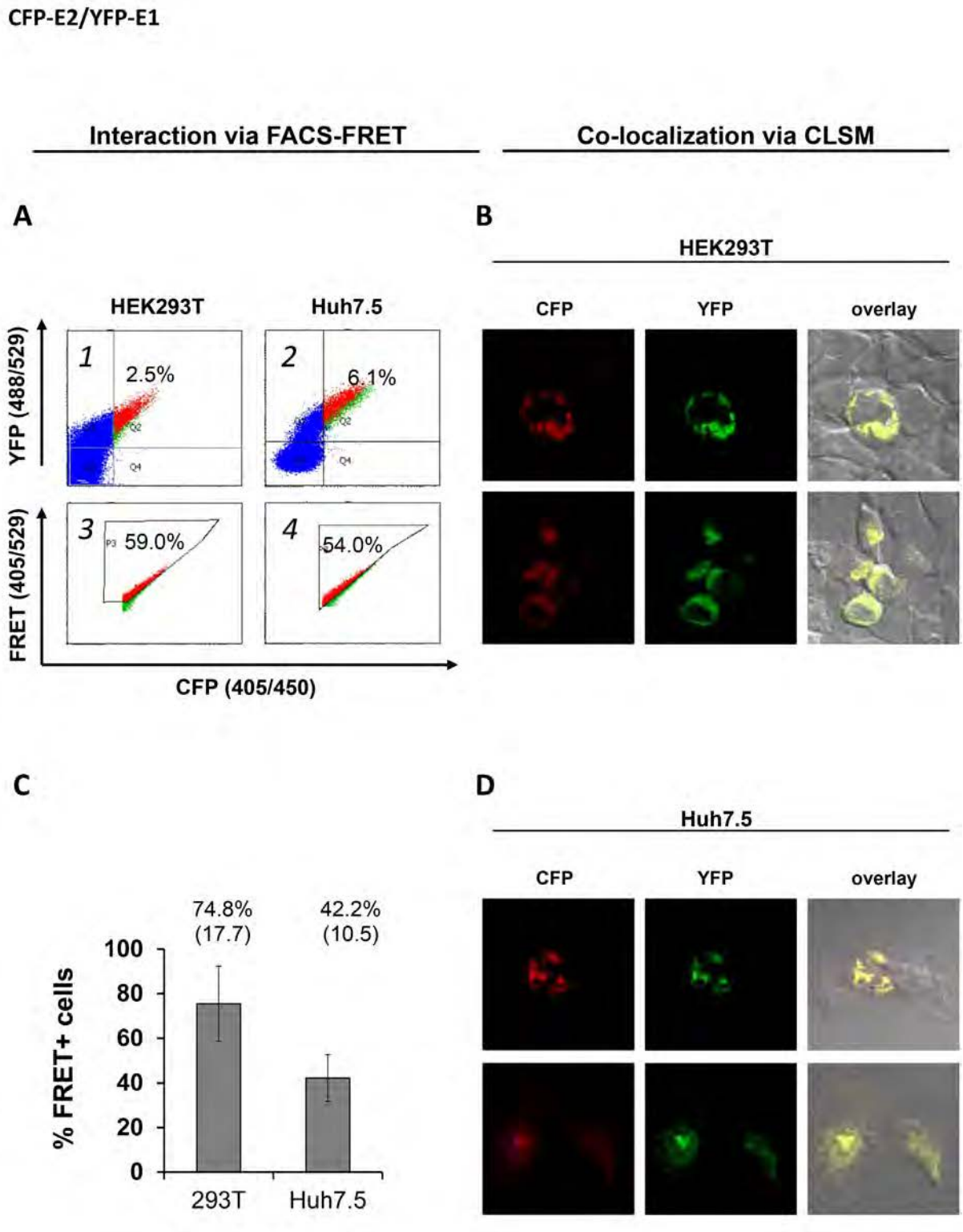

Fig. 27: **E1/E2 complex formation**. Analysis of CFP-E2 and YFP-E1 protein interaction in HEK293T and Huh7.5 cells via FACS FRET. (A) Representative primary FACS plots displaying the amount of co-transfected (*1*, *2*) and FRET-positive cells (*3*, *4*) in both cell lines. Representative confocal images of HEK293T (B) and Huh7.5 (D) cells co-transfected with CFP-E2 (shown in red) and YFP-E1 (green) indicate their co-localization. (C) Mean values and standard deviations (in parentheses) for all performed experiments in HEK293T (N= 17; $p < 0.001$) and Huh7.5 (N= 7; $p < 0.001$) cells confirm multimerization of HCV glycoproteins E1 and E2.

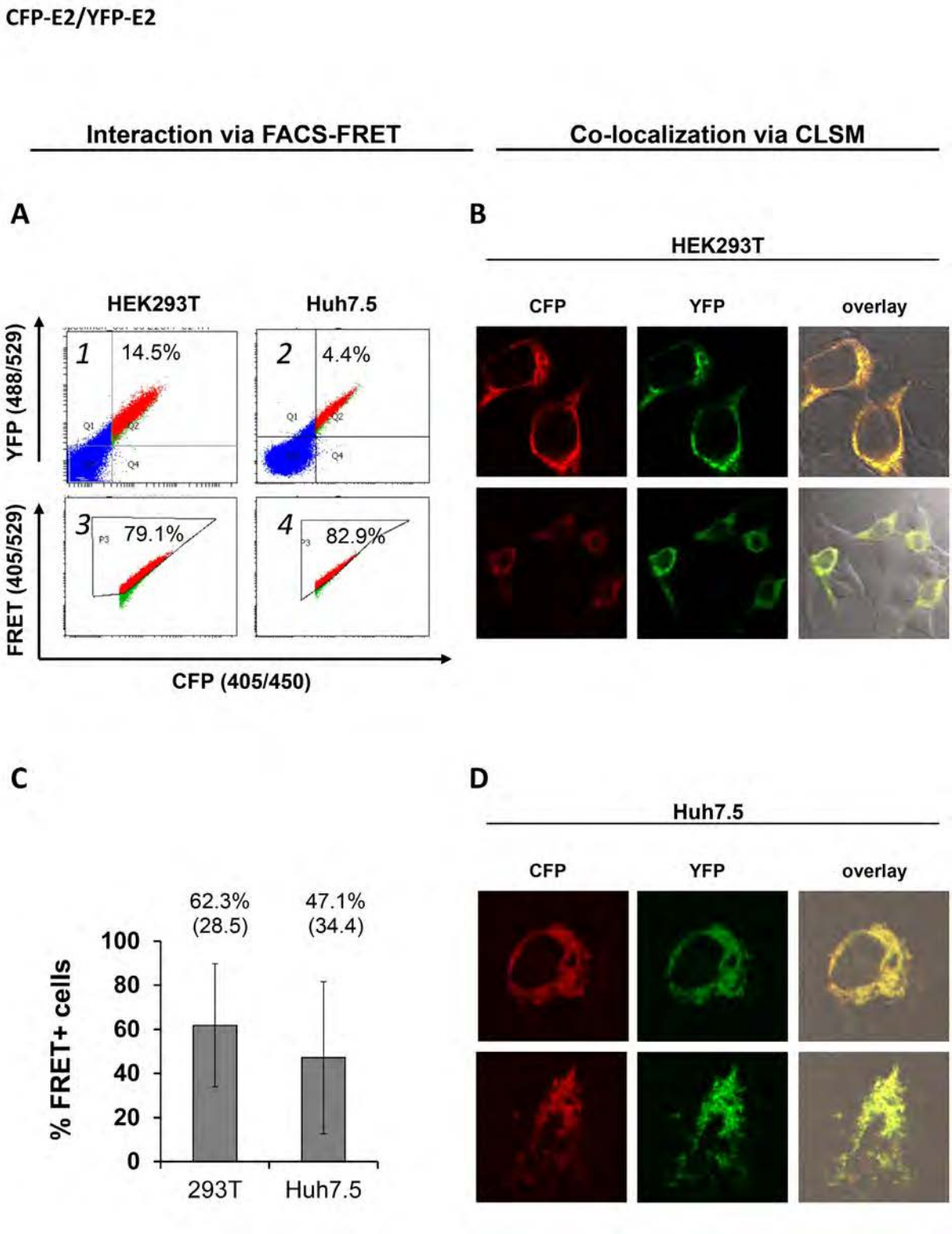

Fig. 28: **E2 multimerization**. Analysis of CFP-E2 and YFP-E2 protein interaction in HEK293T and Huh7.5 cells via flow cytometric energy transfer. (A) Representative primary FACS plots displaying the amount of co-transfected (*1*, *2*) and FRET-positive (*3*, *4*) cells in both cell lines. Confocal images of HEK293T (B) and Huh7.5 (D) cells co-transfected with CFP-E2 (shown in red) and YFP-E2 (green) indicate their co-localization. (C) Mean values and standard deviations (in parentheses) for all performed experiments in HEK293T (N= 18; $p < 0.001$) and Huh7.5 (N= 13; $p < 0.001$) cells confirm homomerization of HCV E2 protein.

We also confirmed the already known interactions of E1 und E2 with NS2 in HEK293T cells (17.7 %, +/- 14.09, n = 34, $p < 0.001$ & 40.1 +/- 32.9, n = 17, $p < 0.001$). In liver cells only the interaction of E2 with NS2 could be verified (16.6 %, +/- 21.98, n = 9, $p < 0.05$).

5.1.5.3 Discovery of Novel Binding Partners by FACS-FRET

Overall, 20 interactions were found within the FACS based FRET approach considering the stringency 10 % threshold (Fig. 27). Eight interactions could only be measured in HEK293Ts, but 12 confirmed in both cell lines. Seven interactions are reported here as novel. These are Core/E2, Core/p7, Core/NS2 and E1/p7 solely in 293T cells (Fig. 29). Interactions, which were observed in both cell lines for the first time, are E2/p7, E1/NS5B and E2/NS5B (Fig. 30 – Fig. 32).

5.1.5.4 HCV Core Interacts with E2

Interaction of Core with E2 was postulated in the literature before. Nevertheless, biochemical approaches failed to detect the interaction thus far. However, interplay of these two HCV proteins is conceivable, since E2 is attached to the capsid. While confirmation of this interaction with biochemical methods such as Co-IP fails, FCET signal in HEK293T cells (21.11 %, +/-17.46, n = 23, $p < 0.001$) was pronounced and statistically highly significant (Fig. 29).

5.1.5.5 Core Interacts with p7 and NS2

It was suggested that p7 is able to regulate the localization of NS2[60], and in turn p7 together with NS2 was postulated to regulate Core localization[9]. This interaction is seen within the FACS based FRET approach in HEK293T (40.73 %, +/-26.82, n = 24) and Huh7.5 cells (14.05 %, +/-11.29, n = 6, $p < 0.01$) as well. In addition, it is demonstrated for the first time that Core interacts directly with both, p7 (27.7, +/-32.91, n = 15, $p < 0.01$) and NS2 (15.5, +/-21.20, n = 35, $p < 0.01$) in living HEK293T cells (Fig. 29).

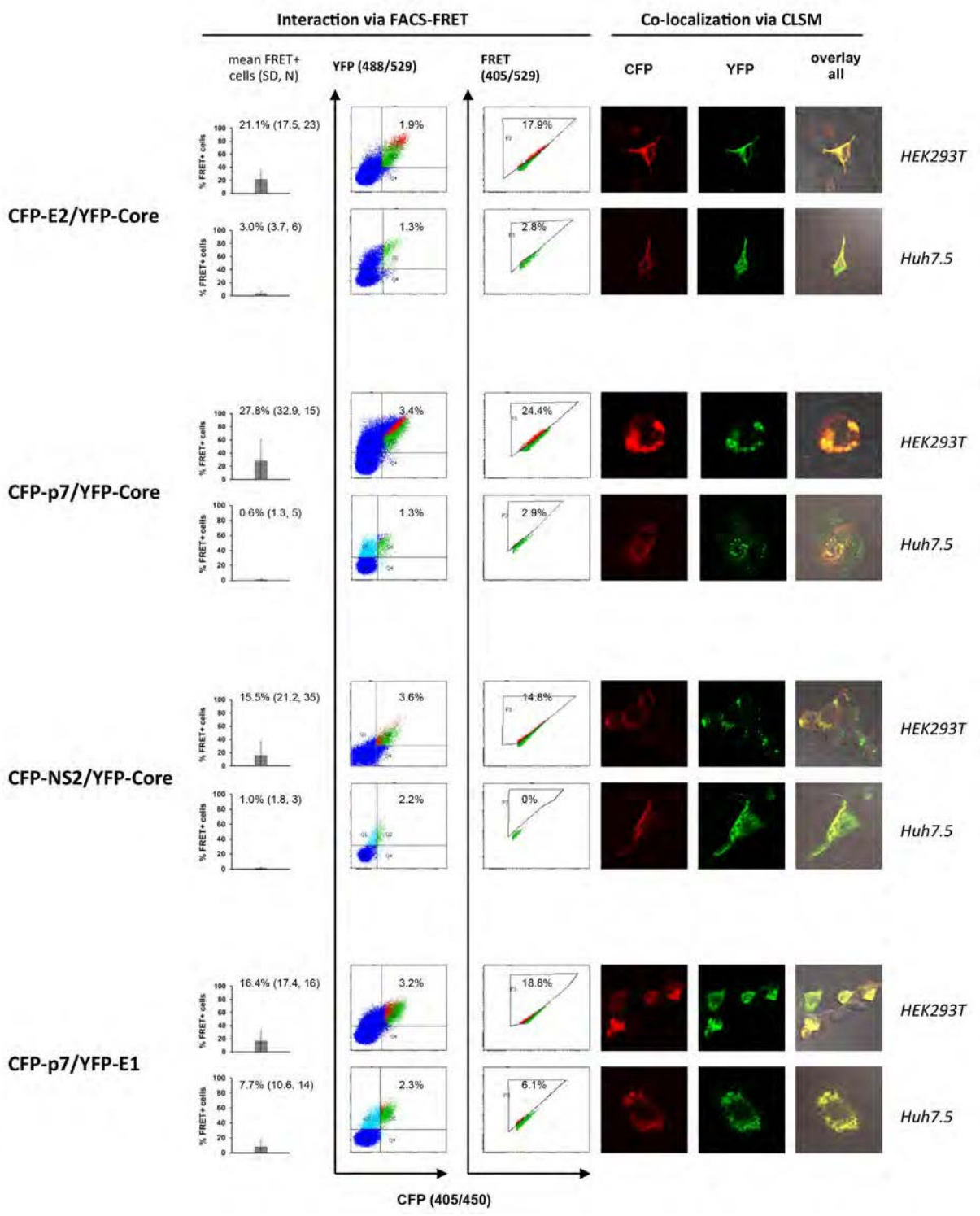

Fig. 29: **New interactions exclusively found in HEK293T cells.** Analysis of mainly YFP-Core interactions with CFP-E2, CFP-p7 and CFP-NS2, and additionally the interaction of CFP-p7 with YFP-E1 in HEK293T and Huh7.5 cells via flow cytometric energy transfer. FRET signal (mean, SD, n) for all performed co-transfections is indicated. Single examples show co-transfection efficiency (%, CFP against YFP, green and red dots) and FRET signal (%, CFP against FRET). FRET-positive cells are indicated by red, FRET-negative cells by green dots. Localization of co-transfected fusions was analyzed via CLSM.

5.1.5.6 Glycoprotein and Viroporin – HCV E2 Interacts with p7

FACS-FRET demonstrates the direct interaction between the two glycoproteins E1 (Fig. 29) and E2 (Fig. 30) with p7 (16.41 %, +/- 17.37, n = 16, $p < 0.001$; 18.36 %, +/- 20.68, n = 17, $p < 0.001$, respectively in HEK293T cells). Of note, regarding the E2/p7 interplay the percentage of FRET-positive Huh7.5 liver cells was twofold higher (34.92 %, +/- 31.22, n = 13, $p < 0.01$) than in HEK293Ts (Fig. 30).

5.1.5.7 NS5B Interacts with Both HCV Glycoproteins

FRET signals with nearly the same intensity in the two tested cell lines were measured for E1 with NS5B (in HEK293T cells: 23.62 %, +/-21.94, n = 20, $p < 0.001$; in Huh7.5 cells 20.55 %, +/- 5.03, n = 4, $p < 0.001$) and E2 with NS5B (in HEK293T cells: 12.97 %; +/- 15.34, n = 18, $p < 0.001$; in Huh7.5 cells: 18.42 %, +/-15.36, n = 6, $p < 0.01$) (Fig. 31 & Fig. 32). This suggests that HCV glycoproteins might regulate polymerase activity; alternatively this interplay could be important for manipulation of host cell signal cascades.

In sum we obtained for a variety of interactions fairly high variations reflected by standard deviations, which were higher than 10 % of the respective mean value. This might be due to the mostly transient nature of the interactions between HCV proteins. However, a large number of biological replicates were done to get statistical confidence for the interactions measured (Fig. 23).

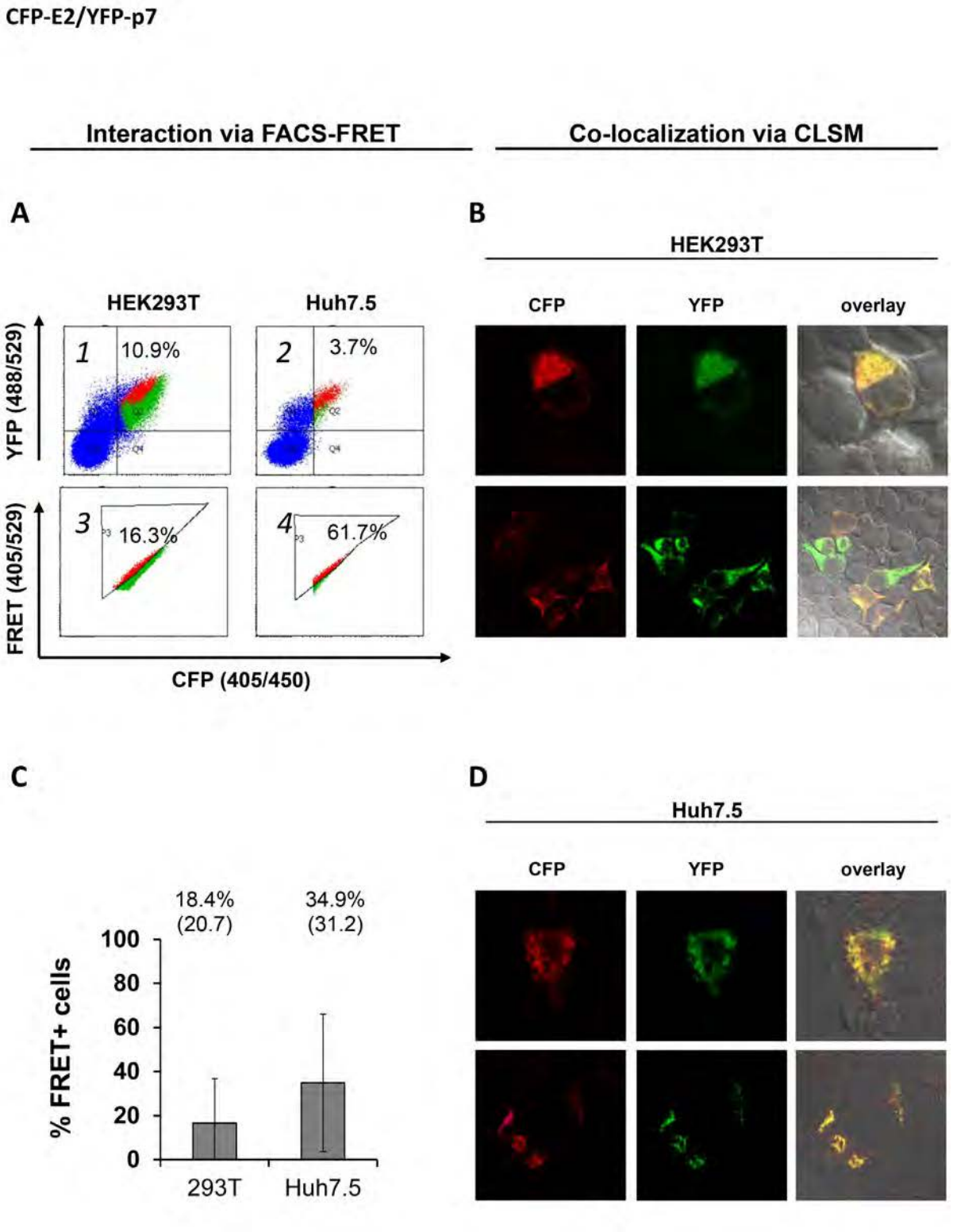

Fig. 30: **E2/p7 interplay**. Analysis of CFP-E2 and YFP-p7 protein interaction in HEK293T and Huh7.5 cells via flow cytometric energy transfer. (A) Representative primary FACS plots displaying the amount of co-transfected (*1*, *2*) and FRET-positive (*3*, *4*) cells in both cell lines. Confocal images of HEK293T (B) and Huh7.5 (D) cells co-transfected with CFP-E2 (shown in red) and YFP-p7 (green) indicate their co-localization. (C) Mean values and standard deviations (in parentheses) for all performed experiments in HEK293T (N= 17, $p < 0.001$) and Huh7.5 (N= 13, $p < 0.01$) cells confirm interaction of HCV glycoprotein E2 with its viroporin p7 in both cell lines.

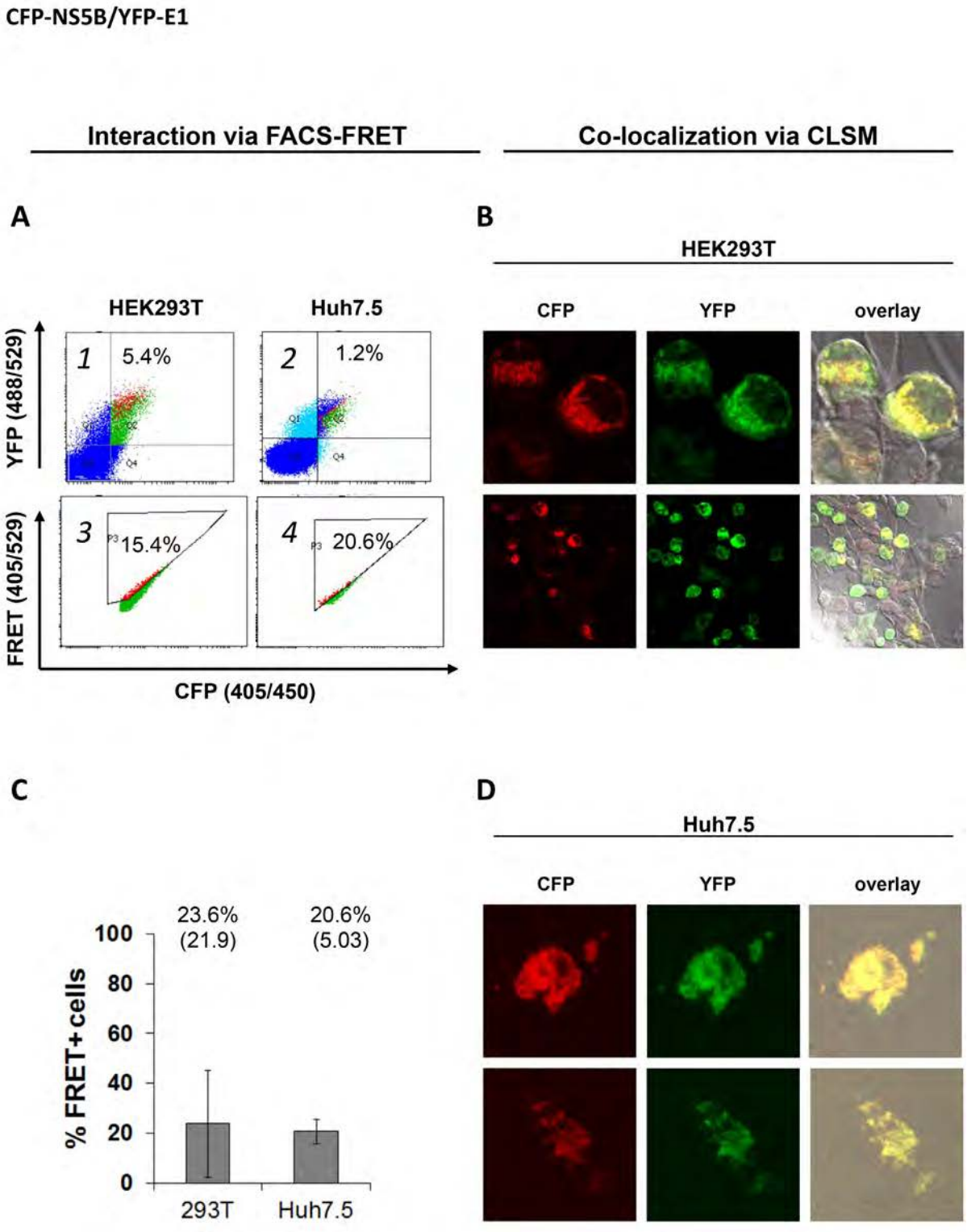

Fig. 31: **NS5B/E1 interaction**. Analysis of CFP-NS5B and YFP-E1 protein interaction in HEK293T and Huh7.5 cells via flow cytometric energy transfer. (A) Representative primary FACS plots displaying the amount of co-transfected (*1*, *2*) and FRET-positive (*3*, *4*) cells in both cell lines. Confocal images of HEK293T (B) and Huh7.5 (D) cells co-transfected with CFP-NS5B (shown in red) and YFP-E1 (green) indicate their co-localization. (C) Mean values and standard deviations (in parentheses) for all performed experiments in HEK293T ($n = 20$, $p < 0.001$) and Huh7.5 ($n = 4$, $p < 0.001$) cells confirm interaction of HCV RNA dependent polymerase NS5B with its glycoprotein E1.

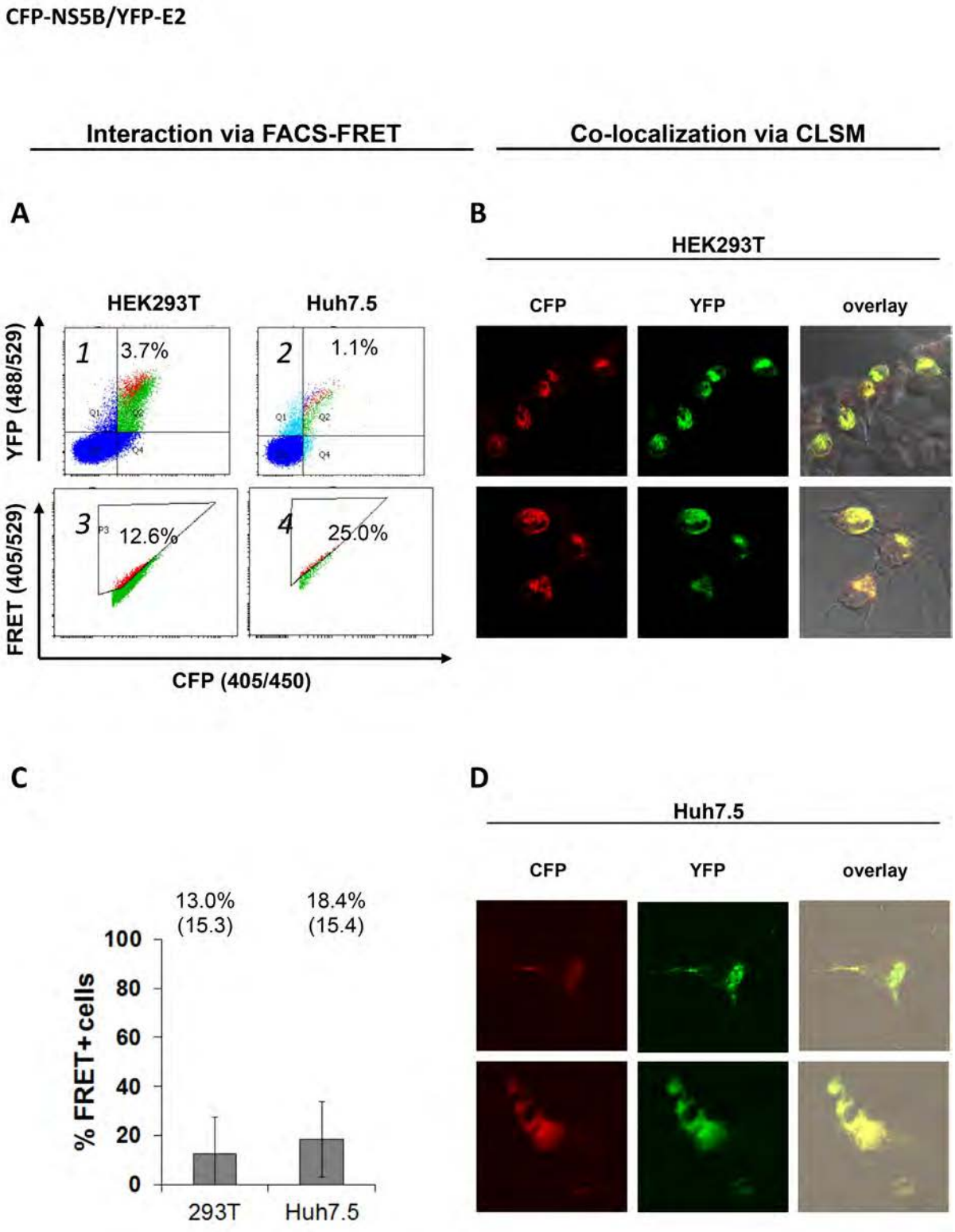

Fig. 32: **NS5B/E2 interaction**. Analysis of CFP-NS5B and YFP-E2 protein interaction in HEK293T and Huh7.5 cells via flow cytometric energy transfer. (A) Representative primary FACS plots displaying the amount of co-transfected (*1*, *2*) and FRET-positive (*3*, *4*) cells in both cell lines. Confocal images of HEK293T (B) and Huh7.5 (D) cells co-transfected with CFP-NS5B (shown in red) and YFP-E2 (green) indicate their co-localization. (C) Mean values and standard deviations (in parentheses) for all performed experiments in HEK293T ($n = 18$, $p < 0.001$) and Huh7.5 ($n = 6$, $p < 0.01$) cells confirm interaction of HCV RNA dependent polymerase NS5B with its glycoprotein E2.

5.1.6 Biochemical Approaches – Co-Immunoprecipitation

Evaluation of found interactions from FCET with co-immunoprecipitation (Co-IP) provides a biochemical approach to confirm reported interactions.

Particularly interesting are the novel interactions Core/E2 and E2/p7. The HCV glycoprotein E2 seems to be associated with Core, which forms the viral nucleocapsid. With FCET this interaction was demonstrated, but never approved before with biochemical methods. In addition, the interaction of E2 with the HCV viroporin p7 could not be detected by IP before. Curiously, the exact function of p7 is not known in detail so far. However, it is supposed to function as a viroporin, shown by Pavlović et al.[20]. Additionally, p7 is responsible for NS2 and Core localization to the ER[9,60] and for assembly and release of infectious virions[154].

5.1.6.1 Co-IP from Transfected HEK293T Cells

The fusion constructs XFP-E2 and XFP-p7 were used for the Co-IP approach in HEK293T cells. Specific antibodies for HCV genotype 2a are only commercially available for Core, E2 and NS5A. Therefore, precipitation of p7 was performed with anti-GFP antibody.

The experimental setup composed controls consisting of different combinations of co-transfections, to exclude false positives due to unspecific binding and allowing precipitation in both directions. Therefore, sepharose-A was incubated either with anti-Core antibody (detection with anti-E2 mAb) or with anti-E2 antibody (detection with anti-Core mAb). For E2/p7, precipitation only in one direction was possible. Due to the missing specific antibody, p7 tracking was achieved with anti-GFP antibody, which detects all transfected fusion proteins.

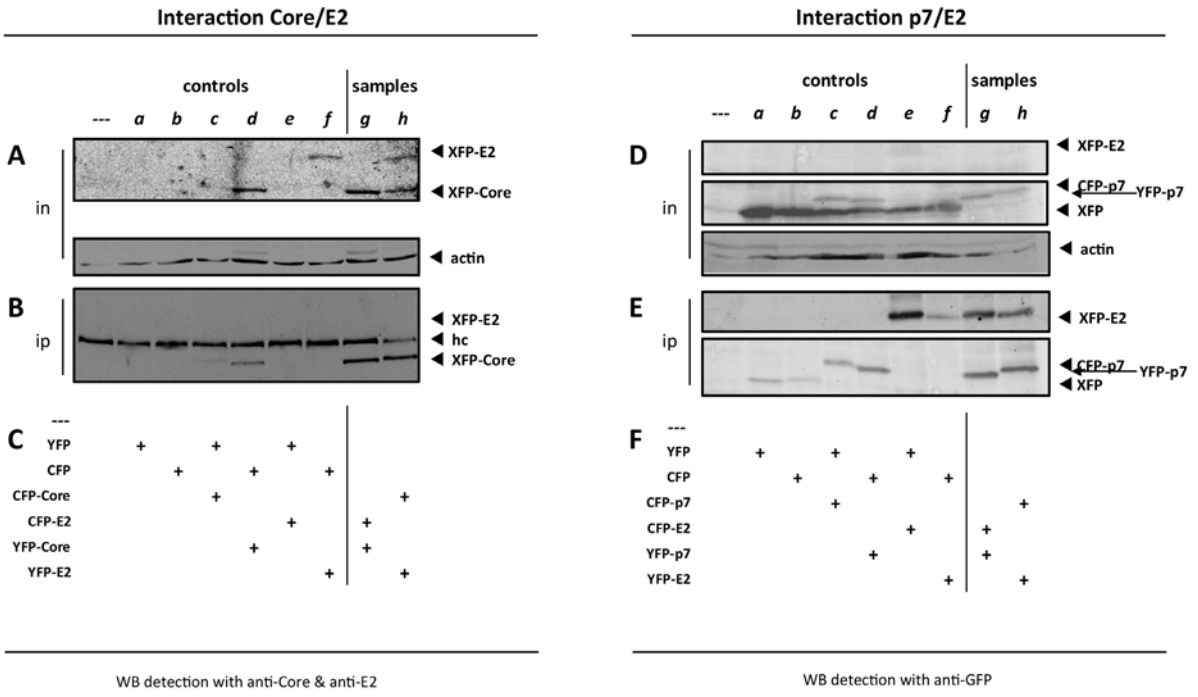

Fig. 33: **Co-IPs from transfected HEK293T cells.** Within the FACS based FRET assay found interactions Core/E2 (A, B, C) and p7/E2 (D, E, F) were tested via co-immunoprecipitation (Co-IP). Depicted are all performed controls (--, *a* to *f*) and actual test-samples (*g* & *h*: fusions co-transfected in both combinations (CFP-protein A + YFP-protein B and YFP-protein A + CFP-protein B) for each investigated interaction). In both cases protein-A sepharose was coupled with anti-HCV-E2 antibody. Detection took place via anti-HCV Core and anti-HCV E2-antibody for Core/E2 interaction (ip; B) and anti-GFP antibody for p7/E2 interaction (ip; E) respectively. The appropriate input is indicated as well (in; A, D); hc equivalent to heavy chain of anti-E2 antibody.

First problems occurred already within the input. It is challenging to get the same transfection efficiency and/or protein expression for the different HCV fusions in parallel samples. These factors always vary strongly and cannot be determined by Bradford measurement, which comprises the whole protein content. Therefore, the transfected proteins were not detectable in all samples and controls via Western-blot, due to low expression levels of the HCV fusions. Hence, it is not advisable to work with different samples for the same Co-IP, since these are not comparable with each other to make an adequate statement.

The second problem showed up while tracking interaction partners of precipitated proteins. Detection of Core from sample (Fig. 33, co-transfection *g* & *h*) is possible, but it can be observed in the controls as well (Fig. 33, co-transfection *c* & *d*).

The same problem occurred for precipitation of p7 with anti-GFP antibody, to show the E2/p7 interplay. CFP and YFP (Fig. 33, *a* – *f*) and samples (Fig. 33, *g* & *h*) came up in the IP. This suggests unspecific binding of the fusions to protein-A sepharose. The large fluorochrome-tag has a size of

about 30 kDa. Therefore, in some cases the tag is larger than the protein itself, and probably interferes with the sepharose.

5.1.6.2 Co-IPs from Infected Huh7.5 Cells

Co-IPs from virally infected/electroporated cells would even more correspond to natural conditions of HCV protein interactions, and avoid working with tagged proteins. Therefore, in the next steps cell lysates from HCV infected Huh7.5 liver cells were used for co-immunoprecipitation analyses. Due to antibody-restriction, only the Core/E2 interaction was examined. Unfortunately, neither binding of Core after E2 precipitation, nor binding of E2 after Core precipitation could be demonstrated (Fig. 34).

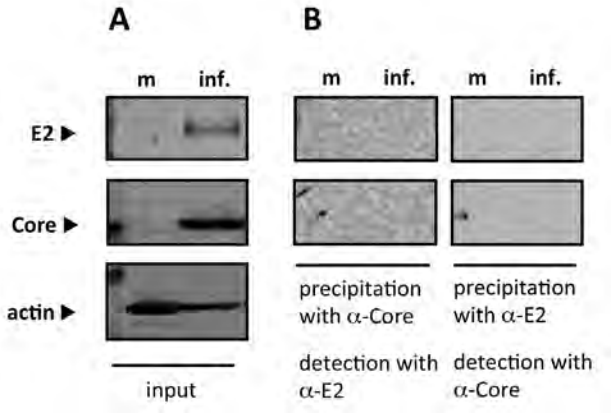

Fig. 34: **Co-IPs from infected Huh7.5 cells.** Whithin the FACS based FRET assay found interaction Core/E2 was tested via co-immunoprecipitation (Co-IP). Depicted are (A) input with mock treated cells (m) and HCV JFH1 infected cells (inf.), as well as (B) precipitation with either α-Core or α-E2 and the corresponding detection of expected binding parters via Western-blot.

5.1.7 HCV Protein-Protein Interaction Network

HCV proteins not only interact with each other but also with host cell proteins. These virus host interactions were not tested in the present approach, but other groups achieved various screens. Nevertheless, no screen for *intra*-viral HCV interactions regarding all HCV proteins was published before. Therefore the FACS based FRET approach has been performed to study novel intra-viral HCV PPIs. To summarize and conclude the found results, an interaction map was constructed with the open source program *Cytoscape*[155]. Intra-viral HCV interactions emerged within the presented FCET approach and intra-viral interactions found within a literature screen (1–48), were summarized in one map (Fig. 35, A). Additional information is provided regarding the interaction of HCV proteins with its host cell proteins. Therefore, interaction data of the VirusMint database (mint.bio.uniroma2.it) was consolidated with data from de Chassey et al. (regarding their S1 supplementary data[156]), who did a proteome-wide interaction screen for HCV. The VirusMint dataset was cleared of double and reverse tested; interactions defined exclusively via co-localization studies were excluded (Fig. 35, B).

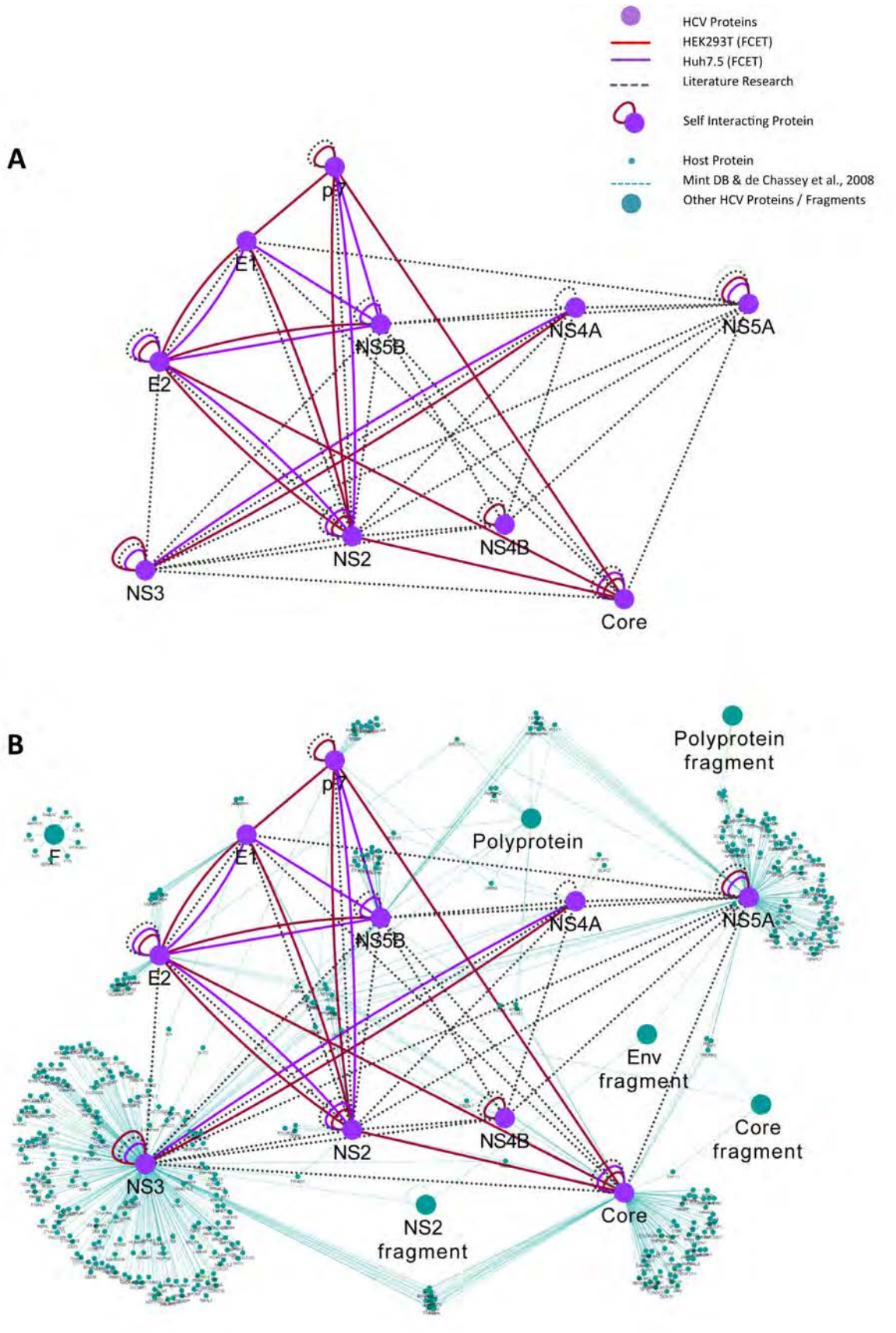

Fig. 35: **Graphical representation of known intra- and inter-viral HCV protein interactions.** Overview of found intra-viral interactions within the FCET approach and literature screen is indicated in A. This network is extended with additional inter-viral interactions of HCV with its host cell proteins in B. Other HCV proteins include fragments of HCV gene products, non-cleaved polyproteins and a frameshift Core-product called F protein. Network was implemented in Cytoscape.

5.2 Expression of E2e

Another intention was to establish an expression system for the HCV glycoproteins with the ultimate goal to solve the structure of HCV E1 and E2. This would be an important step to identify immunogenic epitopes for vaccine development. Furthermore, it will give clues for the generation of novel antivirals.

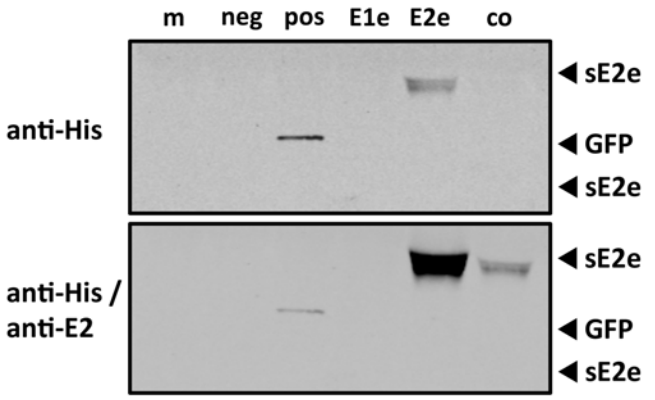

Fig. 36: **Detection of transfected HCV ectodomains** from supernatant of Drosophila Schneider S2 cells via Western-blot. Controls are indicated as m (mock; transfection with water), neg (negative; transfection with pMT/V5/BiP-His) and pos (positive; transfection with pMT/V5-GFP-His). The last three samples indicate single transfection of E1e and E2e and co-transfection of both, which were detected via anti-His and anti-E2 antibody.

Via mass spectroscopy, it was shown that HCV E2 is heavily glycosylated with high-mannose N-glycans[6]. The Drosophila expression system (DES) was chosen, since these cells support protein glycosylation similar to mammalian cells. Furthermore they allow production of high protein amounts sufficient for structural analyses. The expression vector pMT/V5/BiP-His (Invitrogen) contains a BiP secretion signal and a 6X His-tag for protein purification. The ORFs of the HCV glycoprotein ectodomains E1e and E2e were ligated into the vector, checked for correct sequence and transfected via calcium phosphate transfection into Drosophila Schneider S2 cells.

HCV ectodomains were single and co-transfected with calcium phosphate. Co-transfection of the two glycoproteins was performed for a putative subsequent structural analysis of the E1/E2 complex. Furthermore, it was speculated, that E1 and E2 might stabilize each other and give higher protein yields. HCV glycoprotein ectodomain expression is shown in Fig. 36 (m: mock, neg: empty vector, pos: GFP-vector). E1e could not be detected with a His-antibody, or via Coomassie staining (data not shown.

5.2.1 HCV E2e is Functional and Competes with Infectious Virus for Liver Cell Infection

Functionality of a protein strongly indicates that it is correctly folded. Therefore, a competition assay was performed for HCV E2e, which will bind to HCV host cell receptors and therefore inhibit virus binding and entry. Using a reporter virus, the ability of HCV to enter Huh7.5 cells in the presence of E2e was tested. Indeed, the expressed HCV E2 ectodomain from crude supernatant suppresses HCV luc-JC1 infection of liver cells in a dose dependent manner (Fig. 37, A, ◆). To exclude cytotoxic effects by the crude E2e supernatant, an MTT cell cytotoxicity test was performed in parallel. This

experiment revealed that crude supernatant from Schneider insect cells is toxic with increasing volumes (Fig. 37 A, ■).

Thus, the blocking activity of E2e towards cell entry was normalized to the cytotoxicity of the supernatant. This analysis revealed that the E2 ectodomain is properly folded and able to compete with HCV infection of liver cells (Fig. 37, B).

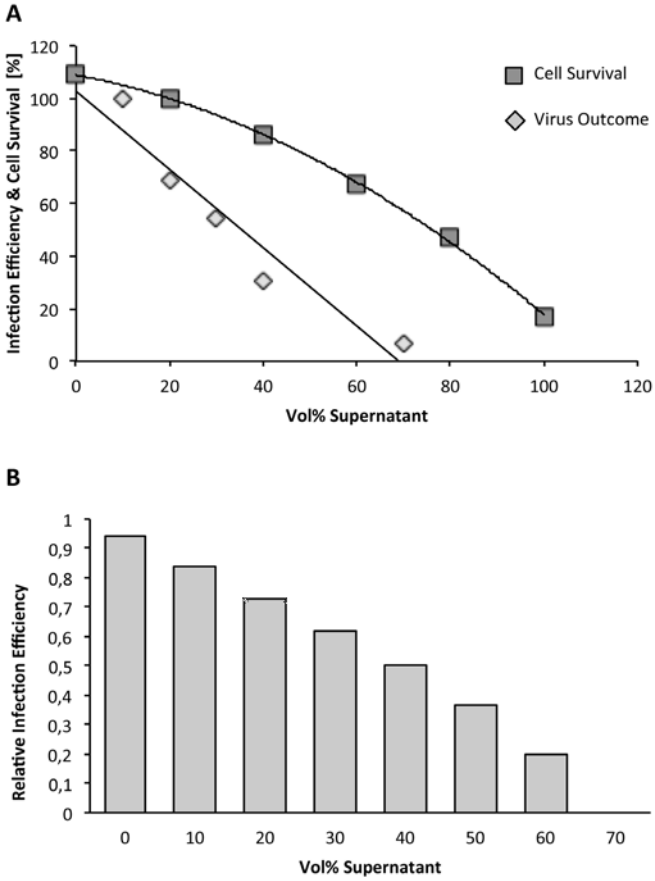

Fig. 37: **Competition Assay.** (A) Huh7.5 cells treated with different amounts (vol%) of E2e containing supernatant. Indicated are mean values of cell survival (MTT test) and HCV infection (luciferase assay). Non-treated cells were set to 100 % of relative cell survival and virus infection. (B) Resulting relative virus infection in competition with E2e. The more supernatant from E2e expressing insect cells, the less luciferase activity could be measured, which indicates inhibition of HCV entry into its host cells.

5.2.2 Ni-NTA Purification

After initial and successful batch purification and its optimization with Ni-NTA sepharose, protein was purified over a Ni-NTA column via the HPLC ÄKTA explorer system (GE-Healthcare). Recovered protein had a higher purity than protein from batch purifications. Quality and purity were estimated via SDS PAGE, followed by Coomassie stain and Western-blot with anti-E2 antibody (Fig. 38). Determination of protein concentration was achieved with Coomassie staining and Bradford Assay (1:100 dilution) in comparison to a BSA standard.

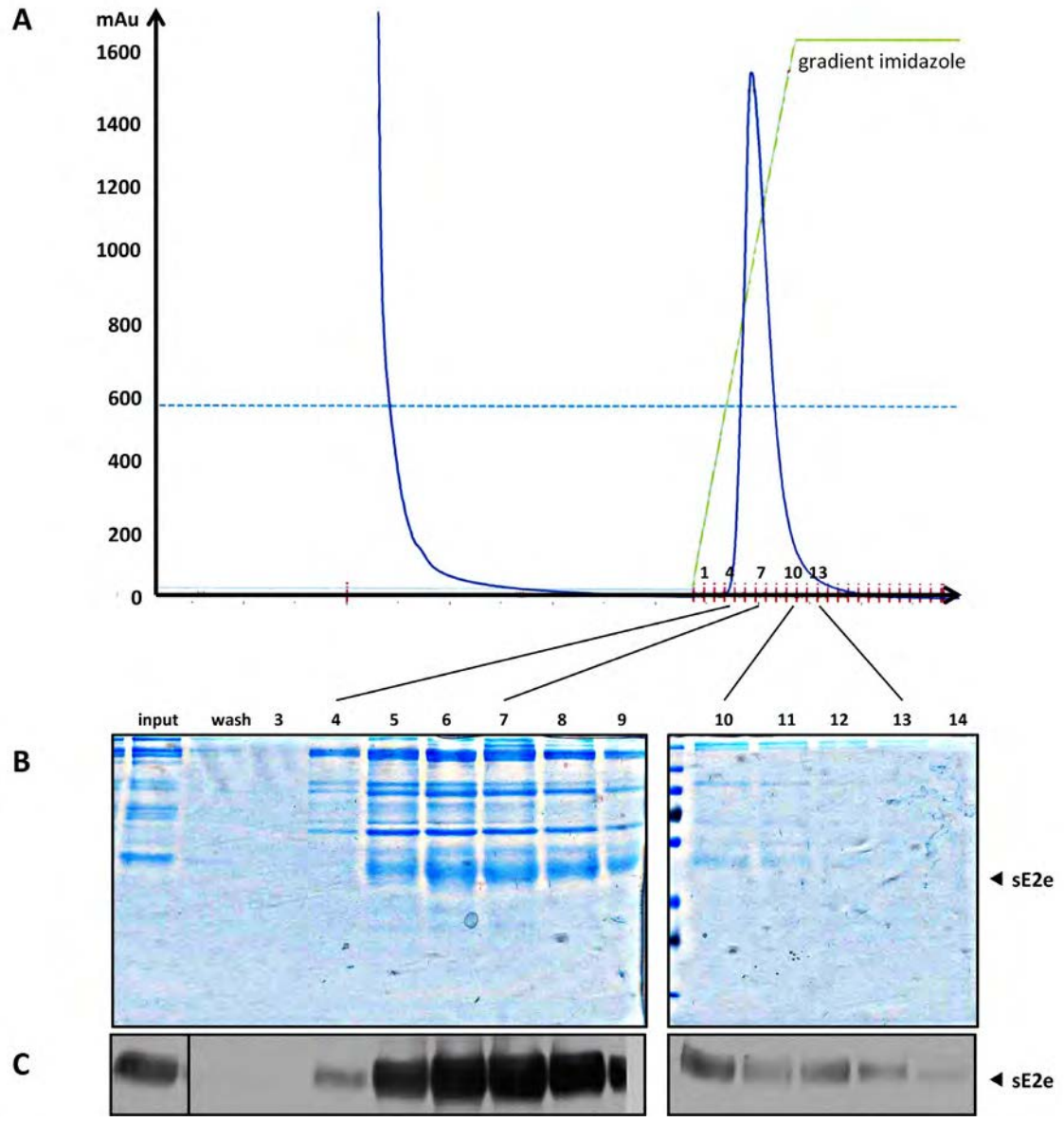

Fig. 38: **Ni-NTA purification of E2e.** (A) Affinity chromatography via Ni-NTA column and HPLC (high performance liquid chromatography). Collected fractions of eluted protein are numbered from 1 to 18. Protein elution was traced via signal measurement at 280 nm during purification. B and C indicate the (B) Coomassie stain of SDS PAGE gel and the corresponding Western-blot (C; ectodomain detection with specific anti-E2 antibody). Input and wash fraction were applied as well. $E2_{661}$ runs with a size of 55 kDa.

5.2.3 Gel filtration

Since additional higher molecular bands could be detected after Ni-NTA purification, gel filtration was exploited as a simple and mild method to separate molecules on the basis of different sizes. We used the Superdex 200 column (GE Healthcare, Fig. 39) and could separate the impurities larger than 200 kDa from E2e, but not the proteins between 60 and 100 kDa. Since separation should be feasible, it is possible that these proteins seen in SDS-PAGE are attached to E2e and separated after SDS treatment and boiling.

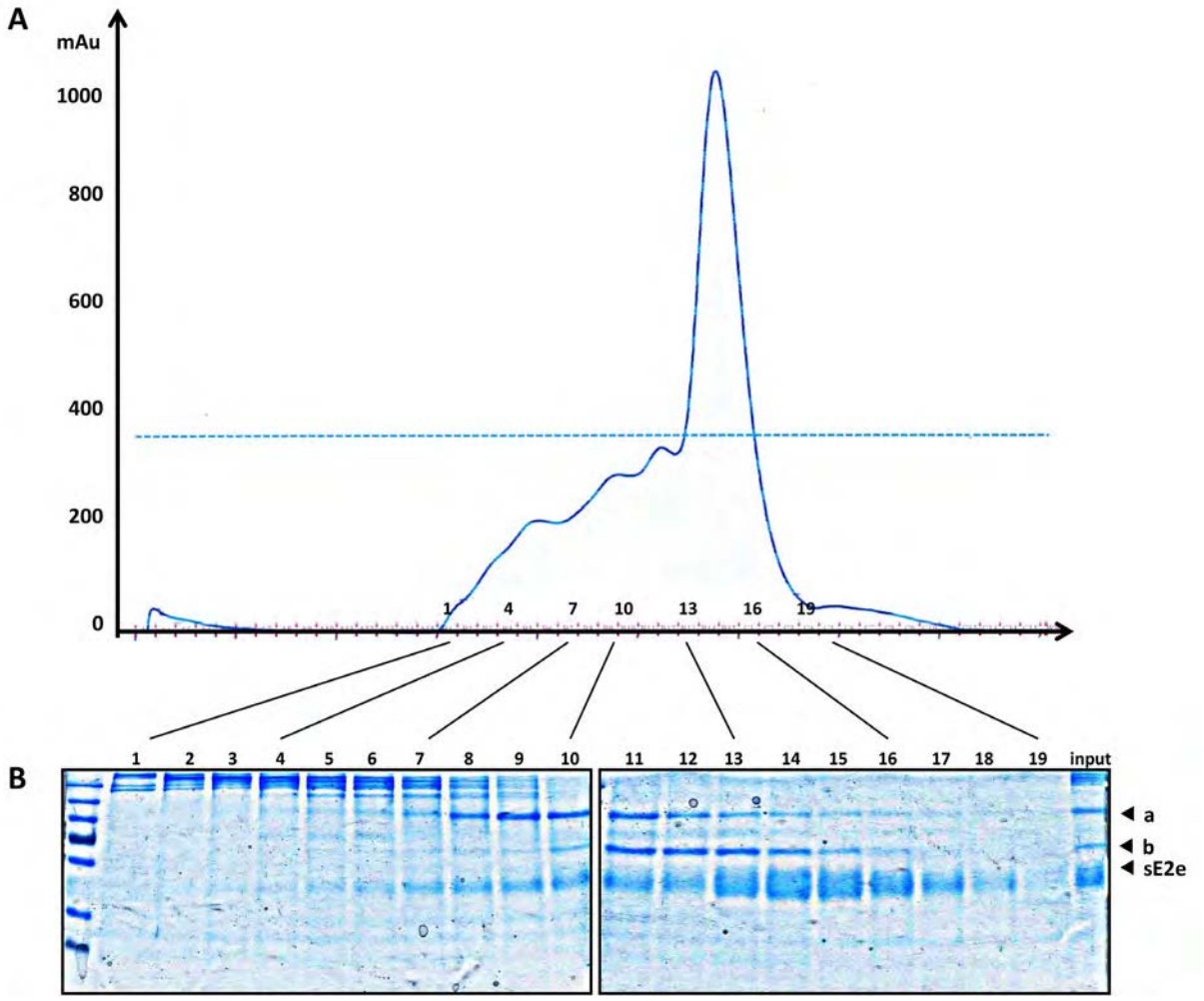

Fig. 39: **Gel-Filtration.** (A) Ni-NTA purified proteins separated via the Superdex 200 column (GE Healthcare). Collected fractions of eluted proteins are numbered from 1 to 19. Protein elution was traced via signal measurement at 280 nm during purification. B indicates the corresponding Coomassie stained SDS-PAGE. Input is shown as well. $E2_{661}$ runs with a size of 55 kDa. Still remaining impurities have a size of approximately 70 (b) and 100 (a) kDa.

5.2.4 Ion Exchange (IEX) – Cation Exchange

Since the theoretical pI of E2e is 8.76 a cation exchange is recommended. However, E2e does not bind to the anion matrix (Fig. 40). Interestingly, the two impurities still seen after gel filtration at 70 and 100 kDa show different behavior regarding the binding affinity to the cation exchanger. Impurity 'a' with a size of about 100 kDa, is binding to the matrix, which is seen for the fractions 40 and 41. Impurity 'b' with a size of about 70 kDa does not bind at all (fraction nb1). This implies different properties regarding the binding affinities of E2e and the impurities 'a' and 'b'. Therefore, a screen has to be performed to find correct conditions for separating these proteins via ion exchange.

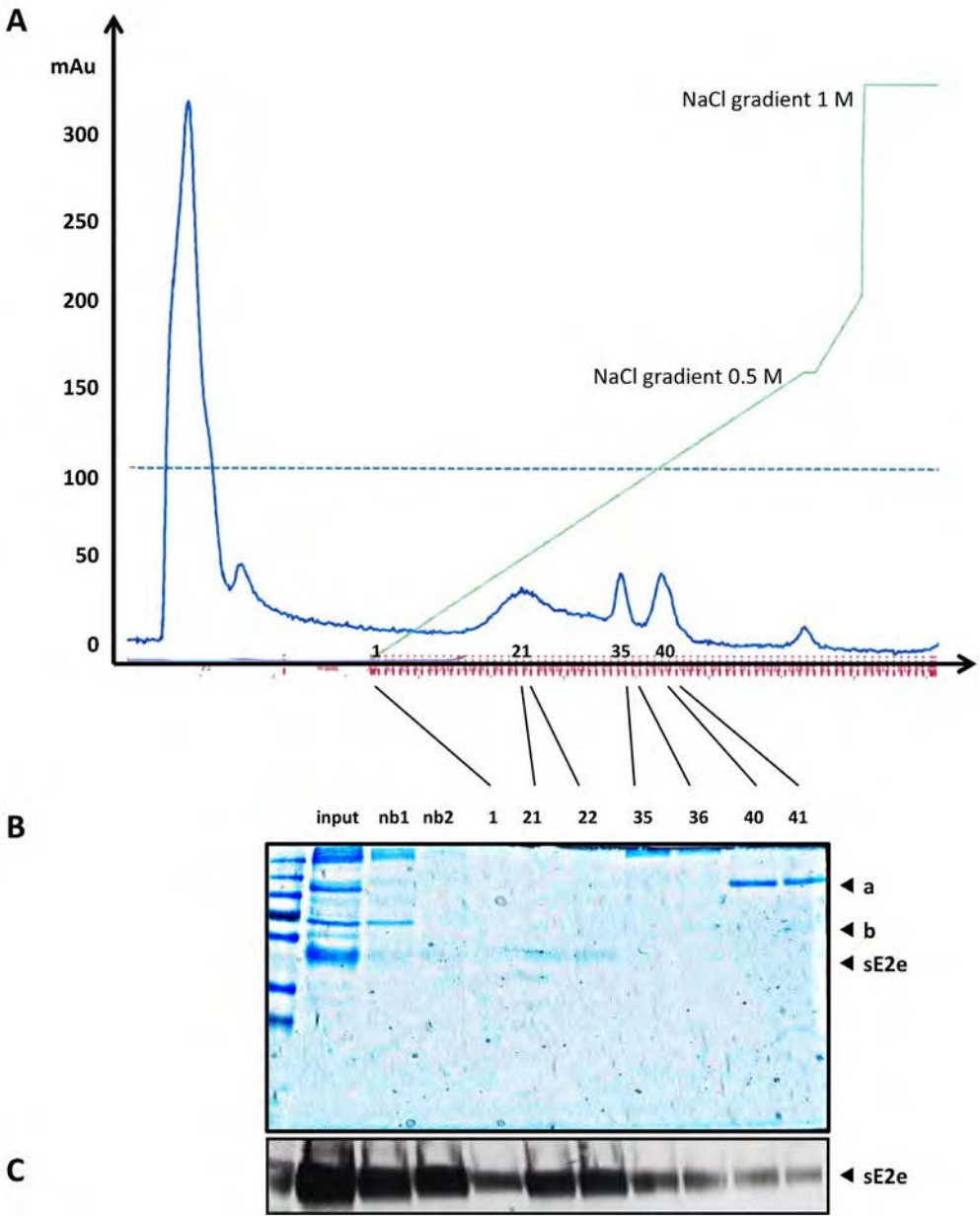

Fig. 40: **Cation exchange with UNO-S (Bio-Rad).** Collected fractions of eluted protein are numbered. Protein elution was traced via signal measurement at 280 nm during separation. B and C indicate the corresponding Coomassie stained SDS-PAGE and Western-blot (ectodomain detection with specific anti-E2 antibody). Input and non-bound fractions (nb) are depicted as well. $E2_{661}$ runs with a size of 55 kDa, impurities 'a' and 'b' have a size of approximately 100 and 70 kDa.

After successful optimization of protein purification, quality control as CD and DLS will be done. Both approaches are a prerequisite for further structural analysis, either X-ray crystallography or SAXS.

In future steps, gained data from SAXS and X-ray experiments can be aligned with structural data from other *Flaviviridae* class-II fusion proteins. As mentioned before, HCV proteins share no sequence similarities with other organisms. However, due to the same function it is likely that the 3D structure of HCV E1 and E2 resembles that from other class-II fusion proteins. Software driven predictions, as given by Krey et al.[153] and Yagnik et al.[5] can be useful to model the structure.

6 Discussion

6.1 FCET

Identification and quantification of protein-protein interactions (PPIs) is essential to understand biological processes. Protein functions in a complex biological system cannot be predicted without any further information. Therefore, it is important to know interaction partners, which theoretically can be predicted with specific tools such as the *OpenPPI Predictor*[157]. To use such programs, orthologous interactome networks from related organisms are needed, as it was shown for *Brugia malayi* using interactome data of *Caenorhabditis elegans* due to a high level of genomic conservation between these two species[157]. Since such orthologous networks are not available for HCV, prediction-based methods are not suitable to generate the intra-viral interaction network of HCV.

Thus, the most reliable way is to look for interactions via experimental measurements. Therefore, a comprehensive and comparative analysis of intra-viral HCV interactions was performed in a medium throughput assay to elucidate the PPI network of HCV. For this purpose, the FACS-based FRET approach was applied, using CFP and YFP-labeled HCV fusions co-transfected into HEK293T and Huh7.5 cells. These cells were analyzed for FRET signals. Furthermore, extensive literature research was performed to align interactions that were previously shown by other groups with the network generated within this thesis.

Post statistical analyses of the FRET results versus background signals and the arbitrary introduction of the stringency threshold of 10 % FRET signal, I identified a total of 20 interactions with FCET. 12 out of these 20 could be verified in Huh7.5 cells and overall seven new interactions of HCV proteins were identified in this thesis: In the HEK293T cell line the interplay of Core/E2, Core/p7, Core/NS2 and E1/p7 was detected. The interactions E2/p7, E1/NS5B and E2/NS5B could be shown in both tested cell lines.

One constraint of this method is the usage of fusion proteins. The functional expression of the fusions used was assessed by extensive quality control. I tested expression levels, transfection efficiency, fluorescence intensity and subcellular localization by single transfections in HEK293T cells. This experiments revealed that expression levels varied, although specific subcellular localization indicated that most fusion proteins are functionally expressed. Western-blot analysis did not give any hint for protein degradation. Nevertheless, HCV proteins with C-terminal chromophore tag were not functional and therefore discarded from further experiments.

For the non-structural proteins many publications with N-terminal-tagged HCV proteins can be found. These proteins were labeled with various tags such as His, FLAG, HA and GST, but no functional constraints were observed,[22,27,48,58,61,65,86] which is in line with the data presented here.

Results from the FCET approach were analyzed with high precaution. Experiments were performed in two cell lines and an extensive number of biological replicates were conducted. Subsequently, statistics were used to assess the significance of identified interactions versus negative controls. Furthermore, an additional stringency threshold of 10 % was introduced. Only FCET results giving FRET signals higher than 10 % were considered as 'true'. Thus, the interactions discussed and presented have a high statistical confidence and were derived from experiments conducted with very high stringency.

6.1.1 Homomerization

In the present work, homomerization was detected for seven out of the ten HCV proteins, exceptions were E1, NS4A and NS5B. This is consistent with already known data regarding homomerization amongst HCV proteins. The role of E1 and its interplay with E2 in this context will be discussed later. NS4A, known as host factor for NS3 protease function, has been shown to be incorporated as monomer in the NS3 complex[28]. This was shown via X-ray crystallography for an NS3 protein complexed with a synthetic NS4A peptide.

The structure of the RNA-dependent RNA polymerase NS5B was also solved via X-ray crystallography by Lesburg et al.[38] for its monomeric state. Nevertheless, Qin et al.[41] showed a strong relationship between oligomerization and activation of NS5B, which possesses at least two critical oligomerization residues. In the FCET approach HCV NS5B displays significant FRET for multimerization, but below the 10 % threshold (5.10 %; +/- 8.72, n = 15). Its structure is mostly studied in combination with inhibitors, but until now the oligomeric state is not clear. Labonte et al.[33] observed that NS5B as an RNA-dependent RNA polymerase is monomeric in solution but is able to build oligomers in presence of short RNAs, which act as template.

For the remaining HCV proteins multimerization could be demonstrated by FCET, in most cases with very high FRET signals. For viral assembly Core multimerization is essential. Core multimerization was shown by various groups[3,49 and others] via crosslinking with glutaraldehyde. In addition, Mousseau et al.[47] demonstrated that Core only interacts with NS3 in its dimeric state. Self-interaction of E2 will be discussed later in more detail. p7 was shown to build an hexameric[16,17,19] or even heptameric[16,57] ion-channel like structure. Its ion-channel function was shown *inter alia* by Pavlović et al.[20] and Griffin et

al.[19] NS2 acts as a dimeric protease, which was demonstrated by Lorenz et al.[62] via X-ray crystallography. Furthermore, NS3 exists as dimer when acting as helicase, which was demonstrated by Cho et al.[65] via X-ray crystallization as well. Oligomerization of NS4B is needed for induction of the membraneous web and additionally NS4B seems to be a key protein for anchorage of other HCV proteins in lipid rafts[158].

NS5A is highly phosphorylated[74] and not much is known about its function. In 2009, Love et al.[159] solved the domain I structure of NS5A via X-ray crystallography and showed that it acts as dimer. No specific enzymatic activity is reported for NS5A, but it is known to affect cellular pathways and innate immune response, host cell growth and stress response. It is postulated that NS5A regulates the switch between HCV replication and assembly[35].

6.1.2 Protein Complexes – The NS3/4A Complex

Taken into account that viral proteins often show multifunctional properties, they can exhibit different functions in a monomeric conformation or in a protein complex. HCV NS3 is one example for a multifunctional protein[113]. Its N-terminus encodes a serine protease, while the C-terminus encodes a RNA helicase/NTPase.

The present work verified the well-known interaction of NS3/NS4A. The N-terminal domain of NS3 encodes a chymotrypsin-like serine protease function. Nevertheless, to be activated NS3 needs NS4A as cofactor to form a stable complex. The NS3/NS4A complex cleaves all proteins downstream of NS3. Co-precipitation studies showed that NS3 and NS4A build a detergent-stable non-covalent complex. The structure of NS3 was determined via X-ray diffraction in presence or absence of NS4A[28,160]. Without its cofactor, the 30 N-terminal NS3 residues are loosely structured. In presence of NS4A, the N-terminal domain forms β-barrels, whereas the structural fold of NS3 at the C-terminus remains unaffected. Thus, NS3 alters its conformation upon binding to its cofactor[161]. It was demonstrated that NS3 expressed alone is mainly found in the cytosolic fraction and proteolytically degraded. This degradation can be abrogated upon co-expression of NS4A, which normally acts as membrane anchor of NS3[162].

Due to the assumption that NS3 expressed alone is degraded and might not be functional, triple transfections were done to stabilize NS3 by co-expression of untagged NS4A. However, FRET signals were unchanged upon triple transfection, indicating that NS3 degradation does not play a major role in the FCET system.

6.1.3 Interplay of HCV Glycoproteins – E1/E2 & E2/E2

HCV E1 and E2 are the viral glycoproteins and mediate entry into the host cell. Oligomerization of viral envelope proteins is essential to control virus assembly and fusion[163]. Stable non-covalent complex formation of both proteins was shown by Deleersnyder et al.[164 and others] Former evidence for a non-monomeric E2 was given by Yagnik et al.[5] They illustrated that E2, a class-II fusion protein, builds a head-to-tail homodimer in heterodimeric association with E1. Additionally the ability of E2e was determined to build mono-, di-, tri- and tetramers[7].

HCV envelope proteins are present in tandem within the Hepatitis polyprotein. Although no amino acid similarities can be observed, class-II fusion proteins show conservation of function and certainly of folding as well[153]. Since viral glycoproteins undergo ER processing before they reach native structure, this may influence each other's folding as well[12]. For *Flaviviridae*, it was shown before, that its viral class-II fusion proteins fold as heterodimer with the glycoprotein encoded upstream in the genome. Membrane fusion is actually driven by dissociation of the complex. For flavivirus prM and alphavirus p62 the first glycoprotein acts as chaperone for the second, while the second acts as membrane fusion protein[165,166]. Brazzoli et al.[12] showed that E1 exists as monomer, but associates with E2 to heterodimers. Therefore, the same mechanism of membrane fusion is assumed for HCV. In the present work, interaction of E1 and E2 (Fig. 27) as well as oligomerization of E2 (Fig. 28) was confirmed via FCET in HEK293T and Huh7.5 cells.

6.1.4 Discovery of Novel Binding Partners by FACS-Based FRET

6.1.4.1 Interplay of the Capsid Core Protein with HCV E2

HCV as enveloped virus needs to anchor E1 and E2 within its capsid, composed of Core proteins. Lo et al.[51] did Co-IPs from a monkey cell line (CV1) and showed the interaction of Core with E1, however, they were not able to detect an interplay of Core and E2. In contrast to these results, interaction of Core with E2 was detected in the present study (Fig. 29), however not with E1. Despite missing FRET signals for the Core/E1 interplay, interaction cannot be excluded.

6.1.4.2 The p7/NS2 Interplay Uncovers New Direct Interactions with Core

The interplay of NS2/p7 has been studied in detail by other groups and could be confirmed using the FCET approach (Fig. 29). Regarding this interaction, interesting observations were done by Boson et al.[9] They verified that p7 and NS2 control the recruitment of Core from the lipid droplets to ER

assembly sites and that Core localization at the ER correlates with the production of infectious particles. Tedbury[60] and others confirmed that NS2 localization depends on p7. This was independent of the p7 ion-channel activity for JFH1 2a sequences.
Popescu et al.[22] reported that Core, the envelope proteins, and p7 influence subcellular NS2 localization. Especially, the interplay of NS2/p7 was confirmed via Co-IP with tagged NS2 and p7 proteins as well as via FRET-FLIM. Ma et al.[58] found NS2/p7 interaction via pull down assay. Furthermore, the interaction of NS2 with NS3 seems to depend on p7[59].

Therefore, NS2 mediates interactions between the replication complex and structural HCV proteins to initiate early steps in virus assembly[59]. It is assumed that NS2 mediates RNA release from the replication complex, for further encapsidation into mature virus particles. The interplay of structural and non-structural proteins for virus assembly appears to be typical for *Flaviviridae*[22].
Using FCET it could be demonstrated for the first time – in accordance with the mentioned results from Boson et al. and Popescu et al. – that there is a direct interaction between p7 and Core as well as between NS2 and Core (Fig. 29). However, other interactions with non-structural proteins could not be confirmed. Interaction with NS proteins may function via a complex or membrane anchorage, which cannot be seen with FCET.

6.1.4.3 HCV E2 Interacts with p7

A typical feature of most animal viruses is to modify host membrane permeability due to virus proteases, glycoproteins or viroporins. Several publications point out that HCV p7 forms a cation selective ion-channel[17,57,167] and therefore acts as viroporin. In general, viroporins are not essential but participate in several viral functions, such as release of viral particle, enhanced viral growth and the passage of ions and small molecules.
Interaction of E2 and p7 was postulated in the literature before[168]. Nevertheless, biochemical approaches failed to detect the interaction so far, most likely due to the transient nature of binding. With FACS-FRET it was possible to demonstrate a direct interaction between E2 and p7 (Fig. 30). Of note, the percentage of FRET positive cells doubled when measurements were done in Huh7.5 cells in comparison to HEK293T cells.

In viruses that lack viroporins, the permeabilization is assumed to be performed by its glycoproteins, as it seems to be the case for HIV-2, which lacks the viroporin Vpu (HIV-1). Both, envelope glycoprotein and viroporin functions are exerted by the HIV-2 Env protein[169]. It is conceivable as well, that both, viral viroporin and glycoproteins could mediate pore formation. Therefore, pore-forming

glycoproteins perhaps accomplish virus entry, in some cases, as well as virus budding. In addition, viroporins are important for virus release[170].

6.1.4.4 E1 & E2 Interact with NS5B

Within the context of this present work, special interest applies to the newly identified interactions, E1/NS5B and E2/NS5B, which appear in both tested cell lines (Fig. 31 & Fig. 32).

In general, HCV polymerase works without glycoproteins as seen for the replicon systems, and as already discussed for its homomerization. However, it is possible that NS5B as well acts as a multifunctional protein, which relies on the interaction with E1 and/or E2. However, the exact natural and biological importance of this interaction needs to be elucidated in more detail using different other approaches. Indeed, interaction of NS5B with the two glycoproteins could also lead to an enhanced activity of RNA-dependent polymerase, or participate in RNA genome incorporation.

6.1.5 Concluding Thoughts Respecting Found Interactions with FCET

With the FCET approach, especially yet not described interactions of structural HCV proteins were newly identified. This is most likely due to the use of replicon systems in the early beginning of HCV research, which encode HCV NS2 or NS3 to NS5B, but not the structural proteins. These replicon systems were able to replicate without structural proteins. Therefore, extensive studies mainly of the HCV NS proteins and their interactions were performed.

For several other *Flaviviridae*, it is known that their replication complexes contain almost all viral proteins[48]. 16 years after HCV discovery, it was possible to also study full-length HCV genomes. Therefore, more and more additional interactions of structural proteins could be elucidated in HCV-infected or electroporated cells. These experiments are based on biochemical methods e.g. Co-IP and will only detect strong physical interactions as well as complexes.

The advantage of the presented FCET system is that all direct intra-viral interactions of HCV proteins could be analyzed, especially in living mammalian cells.

Compared to other viruses, HCV has a very small genome. Human cytomegalovirus (CMV) for example, a DNA virus, possesses 230 kbp with about 170 genes, in contrast to the 3 kbp of HCV. The viruses have to cope with different strategies to enter and to reprogram the host cell, to evade host immunity, to replicate and finally to egress. HCV only has ten proteins to regulate all these processes, which can explain its complex intra-viral interaction network. As indicated before, one possibility for HCV to manage such complex interplay is to build different protein complexes and to use

multifunctional proteins as already described for NS3, which can act as serine protease or as RNA helicase.

Triple transfections were performed to more closely mimic the situation in HCV expressing cells. However, no differences could be observed compared to FRET signals gained from co-transfections, indicating that identified interplays are reliable and only direct interactions are measured. Additionally, in triple complexes of HCV proteins, distances might become too large for energy transfer.

6.1.6 Transient Interactions – Interactions in Living Cells

One important work concerning the PPIs of HCV proteins is the publication of Dimitrova et al.[61] In this work, they used four different approaches to show interactions of the non-structural proteins: GST-pull-down, *in vitro* and *ex vivo* Co-IP as well as Y2H screen.
Dimitrova tested Co-IPs from Huh7 cells transduced with an adenovirus containing HCV NS protein sequences. Many interactions found by Dimitrova, cannot be shown with FCET. Since the HCV replication complex is a large association of various proteins, not all can be linked directly with each other; therefore no FRET can be measured for specific combinations. In contrast to this, indirect interactions can be measured with biochemical approaches. As assembly of HCV takes place at lipid droplets, replication occurs at altered ER membranes in replication complexes, which are associated with microtubules and actin filaments[171]. It is thinkable that association over cellular membrane structures through ER-associated proteins is sufficient to build a stable replication complex of HCV NS proteins.

Additionally, HCV proteins are products of cis- or trans-processing of the polyprotein by viral and cellular proteases. Expressed in an artificial system, mature viral protein likely folds in another way than normal[48], which can explain differences of our results compared to interactions found by other groups.

6.1.7 Differences in Cell Lines

Only homologues of higher primates of CD81 and Occludin and the hepatocyte-specific receptors SR-BI and Claudin 1 mediate entry of HCV independently of the cell type[172]. However, also host factors determine HCV replication and its ability to infect new cells. A recent publication of Da Costa et al.[173] showed that exogenous expression of defined host factors reconstituted the entire life cycle of HCV

in HEK293T cells. The trans-expression of the four HCV entry receptors Occludin, CD81, Claudin 1, SR-BI, of a microRNA abundant in the liver named miR122 mediating HCV replication, and of ApoE an apolipoprotein important for HCV release, support HEK293T cells to enable the complete HCV life cycle. Since HCV production is still diminished in contrast to Huh7.5.1 cells, additional host factors may increase production levels. However, this study implies that important factors for HCV production are generally present in HEK293T cells, and observed differences between the two cell lines tested using the FACS-based FRET approach are most likely due to different expression efficiencies, which explain differences regarding the intra-viral interaction network in liver and kidney cells.

6.1.8 Interplay with Host Proteins

Protein interaction maps to assess the manipulation of the host cell by HCV were generated by various groups. Different strategies yielded a considerable amount of data regarding intra-viral and host-virus protein interactions. Flajolet et al.[48] used an HCV two-hybrid approach including a random genomic HCV library and studied interactions of HCV derived polypeptides and truncated versions. With this approach, they obtained already known and new interactions, e.g. the interaction of NS2 with NS4A, which was additionally confirmed via GST-pull-down.

The first systematic screen of HCV proteins against the human proteome was performed by de Chassey et al.[156], who used a Y2H system to peer for interactions between HCV and human proteins. The gained results were additionally affiliated with extensive literature mining and already known protein interactions from various PPI databases. Thus, they obtained synergy in the context of pathways for insulin, Jak-STAT and TGFβ.

A genome-wide siRNA screen to elucidate host factors for HCV in cell culture was performed by Li et al.[73] with additional bioinformatics meta-analysis, which consolidates experimental results as well as earlier approaches, such as the work of Tai et al.[174] In comparison with Tai et al., 15 out of 96 interactions were detected. Tai identified 96 human genes that support HCV replication within a genome-wide screen. Li and his group ascertained that a significant number (n = 82) of HCV host factors reside in the nucleus. It is noteworthy that only two of the four known host receptors were recovered within this approach (CD81 and Claudin 1). We also detected an interaction between HCV E2 and CD81 using the FCET approach (data not shown).

Different groups investigated the interference of host genes and cellular cofactors[175,176], analyzed interferon-stimulated genes[177], identified human kinases[178], and small molecule regulators[179] which

all decrease the extent of viral replication. Another approach demonstrated that siRNAs directed against HCV proteins can suppress viral replication in a dose-dependent manner[179].

6.1.9 Discussion of FRET

FRET signal intensities are fluctuating to some extent within different combinations, e.g. Core and E2, apparently due to stoichiometric reasons. Especially ECFP-E1 did not show any positive FRET at all. However, four interactions with a FRET signal higher than 10 % were detected for EYFP-E1 in HEK293T cells. Therefore, we circumvent the loss of a specific FRET signal by testing two different combinations. Co-localization levels fluctuate exclusively between neighboring classes. It has to be mentioned here, that visual analysis is not objective and fewer cells can be observed compared to FACS analysis.

As co-localization does not hint directly to interaction, a missing FRET signal is not straightly to interpret as no interaction, since there are some physical limitations regarding energy transfer from one to another protein. However, it is noteworthy for our approach that 'no co-localization' also results in 'no FRET signal' (Fig. 12: C-Core with Y-NS4A, Y-NS4B and Y-NS5A).

6.1.10 Alternative Methods to Detect Protein Interactions

Within Y2H screens, large amounts of sequences can be analyzed with a relatively simple experimental setup. The *in vivo* assay does not require protein purification. However, Y2H screens have a high rate of false positive results. One reason for false positives is that protein expression in yeast does not reflect their expression in the natural environment. Even if proteins interact in yeast, it cannot be verified if they will do so in their natural environment as well, since yeast has a different cellular organization than mammalian cells. Due to the different cellular context or the absence of required post-translational modifications, folding or stability of proteins is likely insufficient. Furthermore, cDNA libraries contain random protein fragments and therefore their interaction depends on specific sub-domains. In reality, interaction of whole proteins can be sterically hindered due to neighboring groups. In addition it is noteworthy, that detected interactions in yeast can be indirect; for instance, endogenous yeast proteins could complex with the analyzed proteins.

Co-IP is a very common and suitable method to precipitate a protein-binding partner from a complex solution, such as cell lysates. A specific antibody is coupled to a solid substrate, like sepharose or magnetic beads. After incubation of cell lysate with an antibody-coupled matrix, proteins are bound to the immobilized antibody. Via Western-blot, additionally bound interaction partners can be

detected. For Co-IPs, proteins can be expressed in their natural cells; therefore, they are in their native state conformation. The main disadvantage is breaking down compartments during cell lysis, which destroys *inter alia* pH and ion concentrations, and therefore the natural context of interaction. Via Co-IP, it is hardly feasible to detect transient interactions and also complexes cannot be discriminated from direct interactions. Furthermore, highly specific antibodies are needed for this method.

The *in vitro* GST-pull-down assays can be done to circumvent the antibody aspect. However, fusion proteins are required. To get high protein amounts, expression often takes place in *E. coli*, which can be problematic for proteins with posttranslational modifications or other features exclusively available in mammalian cells. The GST tag with 26 kDa is quite big in comparison to other tags as FLAG, HA, or c-Myc varying between eight and ten amino acids.

FCET is more suited to detect direct interactions, and important as well to detect these in living mammalian cells. Therefore, the experimental setup is more comparable with the natural environment of examined proteins. Especially cellular factors can be important for full functionality of expressed proteins.

However, we were not able to verify interactions of Core/E2 and E2/p7 detected using FCET biochemically. For the Co-IPs, the same YFP- and CFP-tagged fusions as for the FRET experiments were used. A very high variation of transfection efficiencies was detected. Sometimes the input could not be seen via Western-blot, although fluorescence of cells was checked with a microscope before cell lysis. Additionally, tested fusions interacted unspecific with the precipitation matrix, possibly due to the large YFP/CFP-tag of about 30 kDa. Therefore, fusions with other tags, especially smaller ones such as HA (nine AA) or c-Myc (ten AA) have to be constructed in further experiments to exclude unspecific binding. Another advantage of such tags is that no specific antibodies are needed, and additionally, more interactions found within the FACS-based FRET approach can be examined.

Another approach was to examine intra-viral HCV interactions in the viral context, in parallel circumventing the unspecific binding of tagged proteins. Furthermore, the approach resembles more the natural context of HCV protein expression. Due to the availability of specific antibodies for the HCV 2a genotype, the Core/E2 interaction was tested in infected Huh7.5 cells. No binding of the equivalent protein after precipitation of Core and E2 could be detected. This is in accordance with results of other groups such as Lo et al.[51], which were also not able to verify the interaction of Core

with HCV glycoprotein E2 via precipitation studies. These results are in contrast to our FACS data and knowledge about HCV capsid/envelope organization.

Therefore, to stabilize transient or week interactions, the next step could be to crosslink proteins before precipitation. Crosslinkers with specific linkage sizes are nowadays available to connect proteins at distances not more than the FRET radius. Two examples for membrane soluble crosslinkers are DSP (Dithiobis(succinimidyl propionate), cleavable, 12 Å) and DSS (Disuccinimidyl suberate, non-cleavable, 11.4 Å). Additionally, crosslinkers, which are able to transfer a label tagged with biotin from protein A to protein B upon UV exposure are thinkable. Hereby, the interaction partner of precipitated protein is detected after Western-blot via biotin antibody, labeled avidin or from SDS gel via mass spectroscopy. However, for this approach complex formation has to take place *in vitro*.

Other matrixes than Protein A could also be used for pull down, for example Ni-NTA for poly-His tags (e.g. six-fold His-tag with nine AA). To circumvent the in some cases low and therefore not detectable input, expression can be done in *E. coli*, for example with a GST-tag and glutathione-sepharose as matrix. Here the disadvantages are on the one side the large GST-tag of 26 kDa and on the other site the before mentioned missing posttranslational modifications in the *E. coli* expression system.

It is also possible to work with magnetic bead matrices instead of slurry that can increase purity of precipitated proteins. In addition, complex formation often requires cofactors and energy, which is not given in *in vitro* approaches.

6.1.11 Combination of FRET with a High-Throughput Approach – FACS-Based FRET

In contrast to stable interactions, which are the best-studied ones, transient interplays are short and the most challenging to identify, since the complex may dissociate during the assay. Therefore, FCET is suitable to better understand and substantiate existing data.

Fluorescence signals are much more sensitive than Western-blot signals, which could also be highlighted in the present work for expression levels of CFP-NS4A (Fig. 11). In addition, fluorescent methods expand the field of structural biology. Normally three-dimensional information is acquired via X-ray crystallization and NMR, but fluorescence methods can give additional temporal and spatial information on molecular structures in living cells. Therefore, combination of these methods can result in a more comprehensive picture of biological molecules[141]. FACS-based FRET measurements in

this thesis were performed in accordance to the setup of Banning et al.[90] One advantage of this method is that a heterogeneous mixture and a large number of transfected cells can be studied. This can be done in a short amount of time. Furthermore, compensation can be used to minimize effects of CFP and YFP emission spectral overlap.

FCET is a very strong method to show interactions of proteins, which takes into account to measure *in vivo* interactions with correct subcellular localization of the studied proteins. This method does not disrupt cell compartments, is non-destructive, allows the measurement of direct, but also transient interactions in living cells, and enables to analyze proteins expressed in host cells. The main disadvantage is the need of fusions with YFP and CFP as very large tags (31 kDa). In this context, Siegel et al.[180] pointed out, that FACS-based FRET is a very suitable method to measure small changes in intermolecular distances and that the large tag does not alter the target protein function.

Furthermore, using this technique, some restrictions and physical prerequisites have to be considered:
Since fluorophores can differ in brightness, normally it is recommended to use a FRET pair with the same brightness. Although CFP is fivefold less bright than YFP, the CFP/YFP combination is the best one available for FRET[181], since this pair shows a higher extinction coefficient and quantum yield compared to other fluorochrome combinations, which are available nowadays. Furthermore, CFP has a higher photostability compared to BFP[180].

If two labeled large proteins interact, but the tags are on opposite sites, there is no possibility of energy transfer from one to the other, due to the large distance. In this context, the loss of specific FRET signal has to be accepted and therefore, some interactions probably cannot be shown using this method.

The donor:acceptor dependency is another very important aspect. A ratio ranging outside the 1:10 to 10:1 ratio can limit the FRET signal. One study[182] showed that a 1:2 ratio of donor to acceptor results in a higher FRET signal compared to a 2:1 ratio.
In accordance to that, one interesting observation was done by Koushik et al.[183] In this context, FRET donor (Cerulean) fused to FRET acceptor (Venus) and point mutated amber as non-fluorescent protein, were constructed by the group. The more acceptors were present in the fusion constructs, the higher FRET signals were measured. Additionally, the measured FRET indicated a surplus energy transfer. More energy was measured than expected from summarizing the single energy transfers, which cannot be explained. This fact was also observed in the present work (data not shown).

The FRET phenomenon is known since 1948[139] and nowadays established for various approaches regarding the interaction of biomolecules such as DNA, RNA and proteins. Often FRET is known from fluorescence microscopy, but also from qRT PCR studies. It is even feasible to detect conformational changes in chemical reactions as shown by Maeda et al.[184], where human fucosyltransferase activity is monitored in real time. In this context it is also possible to track conformational changes in protein folding as shown by Kahara et al.[142] FRET is therefore well suited to detect interactions in many different ways. A combination with the high-throughput approach FACS enhances its capacity enormously as demonstrated in the present work.

6.1.12 FACS-Based FRET Approaches are Nowadays Established Methods

To detect PPIs, Kim et al.[185] compared three approaches: FACS-based FRET as it was used in the present work, BRET (bioluminescence resonance energy transfer) and FLIM (fluorescence lifetime imaging microscopy). In that study, the same potential interaction partners of the amyloid precursor protein (APP) were studied using the three different methods.

They proved FRET-based cytometry as the most sensitive and reliable approach to screen for new interactions. In contrast to FCET, FLIM does not depend on fluorophore concentration, but is very tedious. As advantage BRET lacks photo damage and photobleaching. FACS-based FRET was recommended, since it is a non-invasive method to measure large numbers of cells and samples. The derived Z-factor points FCET as very suitable for high throughput-screenings and as an excellent assay for PPIs in comparison with the other tested techniques. FCET provided the most distinct measurement values between interacting and non-interacting proteins and the highest rate of positive hits, without any false positives.

FCET has already been used by our group to discover new and/or confirm already known HIV interactions *in vivo*. Therefore, the interplay of HIV Gag with the host tetraspanin CD81[186], HIV Vpu with host-receptors CD317, CD3, and CD4[90] as well as the interaction of Vpu with host cell Tetherin[187] was approved. Other groups used FCET to show direct interactions of proteins in a similar manner, e.g. Thyrok et al., who showed the interplay of Mint3 and Rab GTPase Rab6A *in vivo*, which was detected before by the group via Y2H and GST-pull-down[188]. Somvanshi et al. demonstrated heterodimerization of cardiac tissue receptors hSSTR5 and β_2AR[189]. Furthermore, the interaction of proteins with specific domains can be highlighted, as it is the case for the recognition of PxxP domains in SAM68 by SH3[190].

A current approach with FACS-based FRET shows that glycosyltransferases form enzymatically active homomeric and heteromeric complexes. Gained results indicated a physical distinction of N-glycosylation and O-glycosylation pathways[191].
Additionally, using FRET-based cytometry it is possible to isolate enzymes from large protein libraries based on their catalytic turn-over[192]. It is also probable to show interactions of RNAs for example to sort for pluripotent stem-cells by binding of specific mRNA donor und acceptor beacons for stem-cell specific transcription factor Oct4 mRNA[193].

In summary, within this thesis the intra HCV protein interaction network in living cells was defined. Already known but also new interactions were described. Additional experiments are needed to define the biological importance of the different interactions. First experiments could define interaction surfaces for example by site-specific mutagenesis or peptide screening. Introduction of these mutations in full-length proviral genomes will give clues on the importance of specific interactions for HCV replication.

6.2 E1 & E2

6.2.1 Characteristics of HCV E1 & E2

The objective of this part of the work was to establish an expression system to produce the HCV glycoproteins E1 and E2. This system should enable three-dimensional structure elucidation of the HCV glycoproteins via X-ray crystallography and small angle X-ray scattering (SAXS) at atomic resolution.

To construct the expression system, the ORFs of the extracellular secreted forms of E1 and E2 were ligated into the respective expression vectors. Secreted forms of viral proteins have been successfully used to elucidate their biological functions and perform structural analyses[144]. Due to the passage across ER and Golgi, only completely post-transcriptionally processed proteins were secreted into the cell culture supernatant, an important point regarding the glycosylation state of HCV E1 and E2. Several groups postulated that these two glycoproteins are interacting with each other. In the present work, FCET verified this circumstance as well. Both proteins contain hydrophobic domains in their C-termini, which act as transmembrane anchors (type-I membrane topology[144]).
Selby et al.[194] showed that the E2 glycoprotein extends to amino acid (AA) residue 746 in the HCV polyprotein. Deletion of the 31 C-terminal AA leads to protein secretion, which is in accordance with Mizushima et al.[195], who decided that the transmembrane domain starts at AA 718.
In the work of Matsuura et al.[15], a truncated E1 which ends at AA position 340 was not secreted. Only when AA 262 up to 290 were deleted in parallel, the protein was secreted into the supernatant. Therefore, this region might acts as a second membrane anchor.

Michalak et al.[144] used Sindbis and vaccinia virus expression systems, in which only truncated E2 is folded and secreted. Truncated $E1_{311}$ was secreted in both systems, but misfolded. The intracellular detection of E1/E2 complexes was possible after co-expression of full-length E2 and truncated E1. Lorent et al.[143] expressed C-terminal truncated $E1_{326}$ in primate kidney cells using recombinant vaccinia virus. They showed that the protein is about 30 kDa in size and can be purified via its C-terminal His tag, which did not affect E1 conformation. For the modeling study of Yagnik et al.[5] (see also introduction) the truncated $E2_{661}$ was used. Importantly, this $E2_{661}$ is sufficient to bind CD81, and is exported and heterodimerizes in a complex with E1.

6.2.2 Expression-Systems for HCV E1 & E2

The type-I transmembrane proteins E1 and E2 consist of a highly glycosylated N-terminal ectodomain, and a C-terminal hydrophobic anchor[13] (Fig. 3 & Fig. 8). These transmembrane domains seem to play a crucial role in both heterodimerization, and subcellular localization[12,52,164]. And although E1 and E2 belong to class-II fusion proteins, they do not have any detectable sequence in common with other members of this protein family. So far, different groups modeled the three-dimensional structure of HCV E2 theoretically, which could be helpful for later interpretation of structural information gained by an X-ray source (X-ray crystallography or SAXS).

In 1994 Matsuura et al.[15], expressed the E1 and E2 proteins in both insect (Baculo system) and mammalian (CHO) cells to gain information about the processing of the individual proteins as well as their interaction with each other. Of special interest for the present work was that E1 and E2 association could still be observed within deletion mutants, lacking internal and C-terminal hydrophobic regions. Matsuura et al. showed that E1 and E2 interact non-covalently, since reducing and non-reducing conditions during SDS-PAGE resulted in no differences regarding the migration pattern.

Yagnik and his group[5] used the TBEV (Tick Borne Encephalitis Virus) envelope protein as template for modeling the HCV E2 structure. Initially they performed various fold recognition methods using software analysis to gain information about the secondary structure of HCV, GBV-A and GBV-B E2, exhibiting genomic sequence similarities. Interestingly, the three proteins showed only a low content of secondary structure (~37 %), which predominantly appeared as β-sheets. The collected data were analyzed further using the prediction program TOPITS (EMBL) that also recognized many β-folds and identified the TBEV E2 protein for similarities in secondary structure despite no sequence similarities. This observation however is fascinating since TBEV and HCV E2 share the same protein function as type-II fusion proteins. Thus, Yagnik et al. proposed a head-to-tail homodimer for HCV E2, as it is the case for TBEV, which would in association with HCV E1, result in a 'homodimeric pair of heterodimers'. After additional mapping of experimental data generated before, they were also able to locate the CD81 and heparin-binding domain in their model.

Folding analyses of E1 and E2 in mammalian cells were performed by Brazzoli et al.[12] It turned out that folding of E1 is faster than that of E2 and that Calnexin is sufficient for E1, but not for E2 folding. In addition the transmembrane-domains of the two glycoproteins are crucial for heterodimer formation and E2 only completes its folding process after association with E1.

Finally, Krey et al.[196] tried to elucidate the 3D structure of E1 and E2. They determined the connectivity of the disulfide bonds and used CD spectroscopy combined with infrared spectroscopy to elucidate the secondary structure of HCV E2, which exhibits 28% of β-sheets. As structural template, they used class-II fusion proteins to model the tertiary organization of the HCV proteins and to map the receptor-binding site for CD81.

Krey et al.[153] provided important information about the predicted three-dimensional structure of the secreted E2 ectodomain (via BiP signal, from an E1E2ΔTMD construct) expressed in Drosophila cells. The isolated E2e reacts with numerous conformation-sensitive mABs and efficiently inhibits HCV-infection of Huh7.5 cells by infectious HCV particles in a dose dependent manner. Therefore, ectodomain expression in insect cells seems to be an adequate approach to get functional proteins for further structural analyses, at least for E2e.

6.2.3 Currently Established Expression System for E2e

To establish an E2e expression system in the present work, a six-fold His-tag was fused C-terminally, replacing the transmembrane domain of the two HCV glycoproteins. An upstream BiP signal ensured E2e secretion into the supernatant after stable transfection. Cells constantly produced a sufficient amount of E2e over a long culture period of several months, and this without any further induction of the expression. Although resistance was encoded on a second co-transfected plasmid, cells did not loose the expression vector over time. In contrast to the pretests in HEK293T cells, where E1e expression has been observed, E1e could neither be detected in the supernatant nor in the cell pellet. Previously, Lorent et al.[143] demonstrated, that HCV E1e is still functional when expressed with a C-terminal His-tag in mammalian and yeast cells. Apparently this is not the case for Drosophila Schneider cells, possibly because expression of E1 has to take place in another system, which also would lead into higher amounts of protein. Another possibility could be to express E1 and E2 in cis instead of in trans.

An initial competition assay showed that expressed E2e is able to compete with HCV JC1 entry, and is therefore suitable for further experiments with respect to later structural analysis.

6.2.4 Protein Purification of E2e

In the next steps E2e protein purification was optimized. After a successful Ni-NTA batch approach, purification was proceeded with Ni-NTA columns and a ÄKTA system, resulting in higher protein

purity. After initial purification with *HisTrap™ FF*, *HisTrap™ Excel* (GE healthcare) Ni-NTA columns were used. These columns are optimized for supernatant purification, since they enable direct loading of unclarified supernatant and exhibit a very strong binding of nickel ions to Ni-Sepharose, which stabilizes their sensitivity against chelating agents in a high amount.
In the higher molecular range some impurities can still be observed via Coomassie staining of SDS-gels, which can be normally eliminated via size exclusion chromatography (SEC) or ion exchange (IEX). SEC is working, but only separates the impurities higher than 200 kDa. However, for a monodisperse solution it is important to get rid of the proteins between 60 and 100 kDa as well.

Due to a theoretical pI of 8.76 a cation exchange was performed to get rid of further impurities. However, cation exchange showed no binding of E2e at all. The isoelectric point of the glycoprotein was calculated based on its amino acid sequence, disregarding charged amino acids eventually shielded by glycosylation, therefore the real pI could be very different to the theoretical. Thus, other conditions are needed for purification via IEX. Interestingly, the impurities showed different binding affinities to the cation exchange matrix than E2e. Component 'a' does bind to the matrix, whereas component 'b' can be retrieved in the non-bound fraction. Therefore, in future steps an ion exchange screen has to be performed to find optimal conditions regarding the protein characteristics, supporting the separation of the two components 'a' and 'b' from E2e.

6.2.5 Future Aspects Regarding the 3D Structure of HCV E1 & E2

With respect to structural analysis of HCV glycoproteins there are still some open questions. One point is, whether folding of both E1 and E2 proteins is dependent on each other, or whether it also occurs separately. Contradictory results regarding this issue were mentioned. Michalak et al.[144] postulated that $E2_{661}$ folding is independent of E1. Brazzoli et al.[12] however, mentioned that E2 completes folding only after association with E1. Since membrane fusion seems to function in a type-II dependent manner where pre- and post-complexes of E1 and E2 can exist, a conformational change from dimer to trimer should be prerequisite for E2 as class-II fusion protein. Another point will be to especially enlighten whether the transmembrane domains are important[163,197] or dispensable for heterodimer building. Possibly the domains assist to bring the proteins into spatial adjacency, but are not primarily responsible for complex formation.

Nuclear magnetic resonance (NMR) is not suited to identify the three-dimensional structure of E2e. This method is only recommended for small proteins up to 35 kDa. For proteins larger than 40 kDa it is challenging to separate distinct signals. Furthermore, due to the isotope labeling needed for NMR,

proteins are often expressed in *E. coli*, which in turn would preclude E2e glycosylation. However, isotopic labeling in S2 cell culture is time-consuming and expensive. Therefore, E2e structure should be elucidated using X-ray crystallography and SAXS. To ensure sufficient quality of the purified protein, this has to be confirmed via circular dichroism (CD) and dynamic light scattering (DLS).
Via CD spectroscopy fractions of α-helices, β-sheets, disordered structures, and disulfide bonds can be determined. DLS, in contrast, measures particle size and protein distribution due to its diffusion coefficient in solution. In general, for a successful crystallization, a monodisperse solution of correct folded, soluble, functional and stable proteins is critical.

After quality control of the E2e protein the crystallization screen can start. Different solutions and/or variation of pH at different temperatures will be tested in 100 nl protein solutions. For this a total amount of five mg pure protein is needed, which makes the optimization of protein purification so important. In parallel SAXS measurements can be performed. In contrast to X-ray crystallography, this method gives lower resolution (10 to 30 Å) and information about the one-dimensional structure. The advantage hereby is, that no protein crystals are needed, and proteins can be analyzed under physiologically conditions. Especially, when proteins contain unstructured parts, which cannot be solved via X-ray crystallography, SAXS data alone but also combined with X-ray data, can give high insights into protein 3D structure.

In general, solving the structure of E1 and E2 can help to substantiate speculations of same functional mechanisms of HCV glycoproteins with glycoproteins of other members of the *Flaviviridae* family and will give more insight into their structural features for vaccine development.

7 Abbreviations

AA	amino acids
BRET	bioluminescence resonance energy transfer
BSA	Bovine Serum Albumin
CD	cluster of differentiation
CFP	Cyan Fluorescence Protein
CHO	Chinese Hamster Ovarian (cells)
CLSM	confocal laser scanning microscopy
Co-IP	co-immunoprecipitation
C-terminal	Carboxy-terminal
DEPC	Diethylpyrocarbonate
DES	Drosophila Expression System
DNA	Deoxyribonucleic acid
ds	double stranded
DTT	Dithiothreitol
E	FRET Efficiency
E. coli	Escherichia coli
ECFP	Enhanced Cyan Fluorescence Protein
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
EtOH	Ethanol
EXFP	enhanced CFP or YFP
EYFP	Enhanced Yellow Fluorescence Protein
FACS	Fluorescence Activated Cell Sorting
FCS	Fetal Calf Serum
FCET	Flow Cytometric Energy Transfer
FLIM	fluorescence lifetime imaging microscopy
FRET	Foersters Resonance Energy Transfer
GBV-A	GB-Virus (first isolated from G. Barker)
GST	glutathione-S-transferase
HBS	HEPES buffered saline
HCV	Hepatitis C Virus
HEK293T	Human Embryonic Kidney (cells)
HEPES	2-(4-(2-Hydroxyethyl)- 1-piperazinyl)-ethansulfonsäure
HIV	Human immunodeficiency virus
HLA	human leucocyte antigen
HPLC	high pressure / performance liquid chromatography
HPV-1	hyper variable region 1
IRES	internal ribosomal entry site
IRF3	interferon regulatory factor 3
Jak	Janus kinase
JFH1	Japanese fulminant hepatitis 1
kbp	kilo base pairs
mAB	monoclonal antibody
MCS	multiple cloning site
MES	2-(N-Morpholino) ethansulfonsäure
MHC	Major Histocompatibility Complex
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazoliumbromid
NEAA	non-essential amino acids

Ni-NTA	Nickel-Nitrilotriacetic acid
NS	non-structural
N-terminal	Amino-terminal
o/n	over night
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline + Tween
PCR	polymerase chain reaction
PFA	Paraformaldehyde
PPI	protein protein interaction
R	Radius
R_0	Foerster Radius
RIG-I	retinoic-acid-inducible gene I
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
rt	real time
SDS	sodium dodecyl sulfate
siRNA	small interfering RNA
SOCS3	suppressor of cytokine signaling
ss	single stranded
STAT3	signal transducer and activator of transcription 3
TBEV	Tick Borne Encephalitis Virus
TBST	Tris Buffered Saline + Tween
TGS	Tris Glycerin Solution
TLR3	toll like receptor 3
UTR	untranslated region
XFP	CFP or YFP
Y2H	Yeast two Hybrid screen
YFP	Yellow Fluorescence Protein

8 References

1. Boulant, S., Vanbelle, C., Ebel, C., Penin, F. & Lavergne, J.-P. Hepatitis C virus Core protein is a dimeric alpha-helical protein exhibiting membrane protein features. J Virol 79, 11353–11365 (2005).
2. Boulant, S. et al. Hepatitis C Virus Core Protein Induces Lipid Droplet Redistribution in a Microtubule- and Dynein-Dependent Manner. Traffic 9, 1268–1282 (2008).
3. Ai, L. S., Lee, Y. W. & Chen, S. S. L. Characterization of Hepatitis C Virus Core Protein Multimerization and Membrane Envelopment: Revelation of a Cascade of Core-Membrane Interactions. J Virol 83, 9923–9939 (2009).
4. Li, H.-F., Huang, C.-H., Ai, L.-S., Chuang, C.-K. & Chen, S. S. L. Mutagenesis of the fusion peptide-like domain of hepatitis C virus E1 glycoprotein: involvement in cell fusion and virus entry. J Biomed Sci 16, 89 (2009).
5. Yagnik, A. T. et al. A model for the hepatitis C virus envelope glycoprotein E2. Proteins 40, 355–366 (2000).
6. Iacob, R. E., Perdivara, I., Przybylski, M. & Tomer, K. B. Mass spectrometric characterization of glycosylation of hepatitis C virus E2 envelope glycoprotein reveals extended microheterogeneity of N-glycans. J. Am. Soc. Mass Spectrom. 19, 428–444 (2008).
7. Rodríguez-Rodríguez, M. et al. Structural properties of the ectodomain of hepatitis C virus E2 envelope protein. Virus Research 139, 91–99 (2009).
8. Bianchi, A., Crotta, S., Brazzoli, M., Foung, S. K. H. & Merola, M. Hepatitis C virus e2 protein ectodomain is essential for assembly of infectious virions. Int J Hepatol 2011, 968161 (2011).
9. Boson, B., Granio, O., Bartenschlager, R. & Cosset, F.-L. A concerted action of hepatitis C virus p7 and nonstructural protein 2 regulates Core localization at the endoplasmic reticulum and virus assembly. PLoS Pathog 7, e1002144 (2011).
10. Suzuki, T. Morphogenesis of infectious hepatitis C virus particles. Front Microbiol 3, 38 (2012).
11. Owsianka, A., Clayton, R. F., Loomis-Price, L. D., McKeating, J. A. & Patel, A. H. Functional analysis of hepatitis C virus E2 glycoproteins and virus-like particles reveals structural dissimilarities between different forms of E2. J Gen Virol 82, 1877–1883 (2001).
12. Brazzoli, M. et al. Folding and dimerization of hepatitis C virus E1 and E2 glycoproteins in stably transfected CHO cells. Virology 332, 438–453 (2005).
13. Lavie, M., Goffard, A. & Dubuisson, J. Assembly of a functional HCV glycoprotein heterodimer. Curr Issues Mol Biol 9, 71–86 (2007).
14. Cocquerel, L. et al. Coexpression of hepatitis C virus envelope proteins E1 and E2 in cis improves the stability of membrane insertion of E2. J Gen Virol 82, 1629–1635 (2001).
15. Matsuura, Y. et al. Processing of E1 and E2 Glycoproteins of Hepatitis C Virus Expresed in Mammalian and Insect Cells. Virology 1–10 (1999).
16. Chandler, D. E., Penin, F., Schulten, K. & Chipot, C. The p7 protein of hepatitis C virus forms structurally plastic, minimalist ion channels. PLoS Comput. Biol. 8, e1002702 (2012).
17. Luik, P. et al. The 3-dimensional structure of a hepatitis C virus p7 ion channel by electron microscopy. Proc Natl Acad Sci USA 106, 12712–12716 (2009).
18. Clarke, D. et al. Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. J Biol Chem 281, 37057–37068 (2006).
19. Griffin, S. D. C. et al. The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. FEBS Lett 535, 34–38 (2003).
20. Pavlović, D. et al. The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. Proc Natl Acad Sci USA 100, 6104–6108 (2003).
21. Welbourn, S. & Pause, A. The hepatitis C virus NS2/3 protease. Curr Issues Mol Biol 9, 63–69 (2007).
22. Popescu, C.-I. et al. NS2 protein of hepatitis C virus interacts with structural and non-structural proteins towards virus assembly. PLoS Pathog 7, e1001278 (2011).
23. Kang, L. W. et al. Crystallization and preliminary X-ray crystallographic analysis of the helicase domain of hepatitis C virus NS3 protein. Acta Cryst (1998). D54, 121-123 [doi:10.1107/S0907444997008883] 1–3 (1998). doi:10.1107/S0907444997008883
24. Cho, H. S. et al. Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. J Biol Chem 273, 15045–15052 (1998).
25. Serebrov, V. & Pyle, A. M. Periodic cycles of RNA unwinding and pausing by hepatitis C virus NS3 helicase. Nature 430, 476–480 (2004).
26. Locatelli, G. A., Spadari, S. & Maga, G. Hepatitis C virus NS3 ATPase/helicase: an ATP switch regulates the cooperativity among the different substrate binding sites. Biochemistry 41, 10332–10342 (2002).
27. Frick, D. N., Rypma, R. S., Lam, A. M. I. & Gu, B. The nonstructural protein 3 protease/helicase requires an intact protease domain to unwind duplex RNA efficiently. J Biol Chem 279, 1269–1280 (2004).
28. Kim, J. L. et al. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. Cell 87, 343–355 (1996).
29. Yao, N. et al. Structure of the hepatitis C virus RNA helicase domain. Nat Struct Biol 4, 463–467 (1997).
30. Tan, S.-L. & Lin, C. HCV NS3-4A Serine Protease. (Horizon Bioscience, 2006).
31. Tan, S.-L., Sklan, E. H. & Glenn, J. S. HCV NS4B: From Obscurity to Central Stage. (Horizon Bioscience, 2006).
32. Welker, M.-W. et al. Dimerization of the hepatitis C virus nonstructural protein 4B depends on the integrity of an aminoterminal basic leucine zipper. Protein Sci. 19, 1327–1336 (2010).
33. Labonte, P. Modulation of Hepatitis C Virus RNA-dependent RNA Polymerase Activity by Structure-based Site-directed Mutagenesis. Journal of Biological Chemistry 277, 38838–38846 (2002).
34. Tellinghuisen, T. L., Foss, K. L. & Treadaway, J. Regulation of Hepatitis C Virion Production via Phosphorylation of the NS5A Protein. PLoS Pathog 4, e1000032 (2008).
35. Tan, S.-L., He, Y., Staschke, K. A. & Tan, S.-L. HCV NS5A: A Multifunctional Regulator of Cellular Pathways and Virus Replication. (Horizon Bioscience, 2006).
36. Macdonald, A. Hepatitis C virus NS5A: tales of a promiscuous protein. Journal of General Virology 85, 2485–2502 (2004).
37. Huang, L. et al. Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. J Biol Chem 280, 36417–36428 (2005).
38. Lesburg, C. A. et al. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Biol 6, 937–943 (1999).
39. Wang, Q. M. et al. Oligomerization and cooperative RNA synthesis activity of hepatitis C virus RNA-dependent RNA polymerase. J Virol 76, 3865–3872 (2002).
40. Tan, S.-L., Ranjith-Kumar, C. T. & Kao, C. C. Biochemical Activities of the HCV NS5B RNA-Dependent RNA Polymerase. (Horizon Bioscience, 2006).
41. Qin, W. et al. Oligomeric interaction of hepatitis C virus NS5B is critical for catalytic activity of RNA-dependent RNA polymerase. J Biol Chem 277, 2132–2137 (2002).
42. Gouklani, H. et al. Hepatitis C virus nonstructural protein 5B is involved in virus morphogenesis. J Virol 86, 5080–5088 (2012).

43. Sillanpää, M. et al. Hepatitis C virus Core, NS3, NS4B and NS5A are the major immunogenic proteins in humoral immunity in chronic HCV infection. Virology Journal 6, 84 (2009).
44. Dancygier, H. Klinische Hepatologie. (Internistische Praxis, 2004). at <http://books.google.de/books?id=iJQZPIAdmpEC&pg=PA413&lpg=PA413&dq=hcv+kda&source=bl&ots=9lJtuvhq1y&sig=mc226a3zQGiKw7WR7eWd9wRYyEE&hl=de&sa=X&ei=2E8iUdiOGszotQaE9oCgCA&ved=0CFcQ6AEwBQ#v=onepage&q=hcv%20kda&f=false>
45. Nielsen, S. U. et al. Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients. J Virol 80, 2418–2428 (2006).
46. Yasui, K. et al. The native form and maturation process of hepatitis C virus Core protein. J Virol 72, 6048–6055 (1998).
47. Mousseau, G., Kota, S., Takahashi, V., Frick, D. N. & Strosberg, A. D. Dimerization-driven interaction of hepatitis C virus Core protein with NS3 helicase. Journal of General Virology 92, 101–111 (2011).
48. Flajolet, M. et al. A genomic approach of the hepatitis C virus generates a protein interaction map. Gene 242, 369–379 (2000).
49. Matsumoto, M., Hwang, S. B., Jeng, K. S., Zhu, N. & Lai, M. M. Homotypic interaction and multimerization of hepatitis C virus Core protein. Virology 218, 43–51 (1996).
50. Merola, M. et al. Folding of hepatitis C virus E1 glycoprotein in a cell-free system. J Virol 75, 11205–11217 (2001).
51. Lo, S. Y., Selby, M. J. & Ou, J. H. Interaction between hepatitis C virus Core protein and E1 envelope protein. J Virol 70, 5177–5182 (1996).
52. Dubuisson, J. et al. Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. J Virol 68, 6147–6160 (1994).
53. Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice, C. M. Expression and identification of hepatitis C virus polyprotein cleavage products. J Virol 67, 1385–1395 (1993).
54. Ralston, R. et al. Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. J Virol 67, 6753–6761 (1993).
55. Lanford, R. E. et al. Analysis of hepatitis C virus capsid, E1, and E2/NS1 proteins expressed in insect cells. Virology 197, 225–235 (1993).
56. Whidby, J. et al. Blocking hepatitis C virus infection with recombinant form of envelope protein 2 ectodomain. J Virol 83, 11078–11089 (2009).
57. Clarke, D. et al. Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. J Biol Chem 281, 37057–37068 (2006).
58. Ma, Y. et al. Hepatitis C virus NS2 protein serves as a scaffold for virus assembly by interacting with both structural and nonstructural proteins. J Virol 85, 86–97 (2011).
59. Stapleford, K. A. & Lindenbach, B. D. Hepatitis C virus NS2 coordinates virus particle assembly through physical interactions with the E1-E2 glycoprotein and NS3-NS4A enzyme complexes. J Virol 85, 1706–1717 (2011).
60. Tedbury, P. et al. The subcellular localization of the hepatitis C virus non-structural protein NS2 is regulated by an ion channel-independent function of the p7 protein. Journal of General Virology 92, 819–830 (2011).
61. Dimitrova, M., Imbert, I., Kieny, M. P. & Schuster, C. Protein-protein interactions between hepatitis C virus nonstructural proteins. J Virol 77, 5401–5414 (2003).
62. Lorenz, I. C., Marcotrigiano, J., Dentzer, T. G. & Rice, C. M. Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. Nature 442, 831–835 (2006).
63. Jones, D. M., Atoom, A. M., Zhang, X., Kottilil, S. & Russell, R. S. A Genetic Interaction between the Core and NS3 Proteins of Hepatitis C Virus Is Essential for Production of Infectious Virus. J Virol 85, 12351–12361 (2011).
64. Schregel, V., Jacobi, S., Penin, F. & Tautz, N. Hepatitis C virus NS2 is a protease stimulated by cofactor domains in NS3. Proc Natl Acad Sci USA 106, 5342–5347 (2009).
65. Cho, H. S. et al. Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. J Biol Chem 273, 15045–15052 (1998).
66. Gallinari, P. et al. Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus. J Virol 72, 6758–6769 (1998).
67. Bartenschlager, R., Lohmann, V., Wilkinson, T. & Koch, J. O. Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. J Virol 69, 7519–7528 (1995).
68. Failla, C., Tomei, L. & De Francesco, R. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. J Virol 68, 3753–3760 (1994).
69. Lin, C., Prágai, B. M., Grakoui, A., Xu, J. & Rice, C. M. Hepatitis C virus NS3 serine proteinase: trans-cleavage requirements and processing kinetics. J Virol 68, 8147–8157 (1994).
70. Tanji, Y., Hijikata, M., Satoh, S., Kaneko, T. & Shimotohno, K. Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. J Virol 69, 1575–1581 (1995).
71. Paredes, A. M. & Blight, K. J. A genetic interaction between hepatitis C virus NS4B and NS3 is important for RNA replication. J Virol 82, 10671–10683 (2008).
72. Lin, C., Wu, J. W., Hsiao, K. & Su, M. S. The hepatitis C virus NS4A protein: interactions with the NS4B and NS5A proteins. J Virol 71, 6465–6471 (1997).
73. Li, Q. et al. A genome-wide genetic screen for host factors required for hepatitis C virus propagation. Proc Natl Acad Sci USA 106, 16410–16415 (2009).
74. Yu, G.-Y., Lee, K.-J., Gao, L. & Lai, M. M. C. Palmitoylation and polymerization of hepatitis C virus NS4B protein. J Virol 80, 6013–6023 (2006).
75. Masaki, T. et al. Interaction of hepatitis C virus nonstructural protein 5A with Core protein is critical for the production of infectious virus particles. J Virol 82, 7964–7976 (2008).
76. Miyanari, Y. et al. The lipid droplet is an important organelle for hepatitis C virus production. Nat Cell Biol 9, 1089–1097 (2007).
77. Goh, P. Y. et al. The hepatitis C virus Core protein interacts with NS5A and activates its caspase-mediated proteolytic cleavage. Virology 290, 224–236 (2001).
78. Lundin, M., Lindström, H., Grönwall, C. & Persson, M. A. A. Dual topology of the processed hepatitis C virus protein NS4B is influenced by the NS5A protein. J Gen Virol 87, 3263–3272 (2006).
79. Love, R. A., Brodsky, O., Hickey, M. J., Wells, P. A. & Cronin, C. N. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. J Virol 83, 4395–4403 (2009).
80. Tellinghuisen, T. L., Marcotrigiano, J. & Rice, C. M. Structure of the zinc-binding domain of an essential component of the hepatitis C virus replicase. Nat Cell Biol 435, 374–379 (2005).
81. Uchida, M. et al. Hepatitis C virus Core protein binds to a C-terminal region of NS5B RNA polymerase. Hepatol Res 22, 297–306 (2002).
82. Kang, S.-M. et al. Regulation of hepatitis C virus replication by the Core protein through its interaction with viral RNA polymerase. Biochem Biophys Res Commun 386, 55–59 (2009).
83. Zhang, C. et al. Stimulation of hepatitis C virus (HCV) nonstructural protein 3 (NS3) helicase activity by the NS3 protease domain and by HCV RNA-dependent RNA polymerase. J Virol 79, 8687–8697 (2005).
84. Ishido, S., Fujita, T. & Hotta, H. Complex formation of NS5B with NS3 and NS4A proteins of hepatitis C virus. Biochem Biophys Res Commun 244, 35–40 (1998).
85. Shirota, Y. et al. Hepatitis C virus (HCV) NS5A binds RNA-dependent RNA polymerase (RdRP) NS5B and modulates RNA-dependent RNA polymerase activity. J Biol Chem 277, 11149–11155 (2002).
86. Shimakami, T. et al. Effect of interaction between hepatitis C virus NS5A and NS5B on hepatitis C virus RNA replication with the hepatitis C virus replicon. J Virol 78, 2738–2748 (2004).

87. Choo, Q. L. et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244, 359–362 (1989).
88. Lindenbach, B. D. Complete Replication of Hepatitis C Virus in Cell Culture. Science 309, 623–626 (2005).
89. Wakita, T. et al. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. Nature Medicine 11, 791–796 (2005).
90. Banning, C. et al. A flow cytometry-based FRET assay to identify and analyse protein-protein interactions in living cells. PLoS ONE 5, e9344 (2010).
91. Phizicky, E. M. & Fields, S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev 59, 94–123 (1995).
92. Klotz, I. M., Langerman, N. R. & Darnall, D. W. Quaternary structure of proteins. Annu Rev Biochem 39, 25–62 (1970).
93. Berggård, T., Linse, S. & James, P. Methods for the detection and analysis of protein–protein interactions. Proteomics 7, 2833–2842 (2007).
94. Stelzl, U. et al. A human protein-protein interaction network: a resource for annotating the proteome. Cell 122, 957–968 (2005).
95. Tarassov, K. et al. An in vivo map of the yeast protein interactome. Science 320, 1465–1470 (2008).
96. Ideker, T. & Krogan, N. J. Differential network biology. Molecular Systems Biology 8, 1–9 (2012).
97. Bailey, J. An assessment of the role of chimpanzees in AIDS vaccine research. Altern Lab Anim 36, 381–428 (2008).
98. Kim, J. W. & Wang, X. W. Gene expression profiling of preneoplastic liver disease and liver cancer: a new era for improved early detection and treatment of these deadly diseases? Carcinogenesis 24, 363–369 (2003).
99. Bailey, J. An assessment of the use of chimpanzees in hepatitis C research past, present and future: 2. Alternative replacement methods. Altern Lab Anim 38, 471–494 (2010).
100. Marcellin, P. Hepatitis B and hepatitis C in 2009. Liver Int 29 Suppl 1, 1–8 (2009).
101. Brown, R. S. Hepatitis C and liver transplantation. Nat Cell Biol 436, 973–978 (2005).
102. Roche, B. & Samuel, D. Hepatitis C virus treatment pre- and post-liver transplantation. Liver Int 32 Suppl 1, 120–128 (2012).
103. Castera, L. Transient elastography and other noninvasive tests to assess hepatic fibrosis in patients with viral hepatitis. J Viral Hepat 16, 300–314 (2009).
104. Tohme, R. A. & Holmberg, S. D. Is sexual contact a major mode of hepatitis C virus transmission? Hepatology 52, 1497–1505 (2010).
105. Gravitz, L. Introduction: a smouldering public-health crisis. Nature 474, S2–4 (2011).
106. Eisenstein, M. Vaccines: a moving target. Nature 474, S16–7 (2011).
107. Ly, K. N. et al. The increasing burden of mortality from viral hepatitis in the United States between 1999 and 2007. Ann. Intern. Med. 156, 271–278 (2012).
108. Operskalski, E. A. & Kovacs, A. HIV/HCV co-infection: pathogenesis, clinical complications, treatment, and new therapeutic technologies. Curr HIV/AIDS Rep 8, 12–22 (2011).
109. Alter, M. J. Epidemiology of viral hepatitis and HIV co-infection. JOURNAL OF HEPATOLOGY 44, S6–9 (2006).
110. Fontaine, H. & Pol, S. Antiviral activity of telaprevir and boceprevir for the treatment of hepatitis C virus infection in treatment-experienced patients. Clin Res Hepatol Gastroenterol 35 Suppl 2, S59–63 (2011).
111. Asselah, T. & Marcellin, P. Direct acting antivirals for the treatment of chronic hepatitis C: one pill a day for tomorrow. Liver Int 32 Suppl 1, 88–102 (2012).
112. Kronenberger, B. & Zeuzem, S. New developments in HCV therapy. J Viral Hepat 19 Suppl 1, 48–51 (2012).
113. Moradpour, D., Penin, F. & Rice, C. M. Replication of hepatitis C virus. Nat Rev Microbiol 5, 453–463 (2007).
114. Novoa, R. R. et al. Virus factories: associations of cell organelles for viral replication and morphogenesis. Biol. Cell 97, 147–172 (2005).
115. Samsa, M. M. et al. Dengue Virus Capsid Protein Usurps Lipid Droplets for Viral Particle Formation. PLoS Pathog 5, e1000632 (2009).
116. Cheung, W. et al. Rotaviruses Associate with Cellular Lipid Droplet Components To Replicate in Viroplasms, and Compounds Disrupting or Blocking Lipid Droplets Inhibit Viroplasm Formation and Viral Replication. J Virol 84, 6782–6798 (2010).
117. Laskus, T. et al. Negative-strand hepatitis C virus (HCV) RNA in peripheral blood mononuclear cells from anti-HCV-positive/HIV-infected women. J Infect Dis 195, 124–133 (2007).
118. Fletcher, N. F. & McKeating, J. A. Hepatitis C virus and the brain. Journal of Viral Hepatitis 19, 301–306 (2012).
119. Ye, J. Reliance of host cholesterol metabolic pathways for the life cycle of hepatitis C virus. PLoS Pathog 3, e108 (2007).
120. Ye, J. Hepatitis C virus: a new class of virus associated with particles derived from very low-density lipoproteins. Arteriosclerosis, Thrombosis, and Vascular Biology 32, 1099–1103 (2012).
121. Korzaya, L. I., Lapin, B. A., Keburiya, V. V. & Chikobava, M. G. Spontaneous infection of lower primates with hepatitis C virus. Bull. Exp. Biol. Med. 133, 178–181 (2002).
122. Dorner, M. et al. A genetically humanized mouse model for hepatitis C virus infection. Nature 474, 208–211 (2011).
123. Lohmann, V. et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. Science 285, 110–113 (1999).
124. Kamrud, K. I., Olson, K. E., Higgs, S., Carlson, J. O. & Beaty, B. J. Use of the Sindbis replicon system for expression of LaCrosse virus envelope proteins in mosquito cells. Arch Virol 143, 1365–1377 (1998).
125. Naito, T. et al. An influenza virus replicon system in yeast identified Tat-SF1 as a stimulatory host factor for viral RNA synthesis. Proc Natl Acad Sci USA 104, 18235–18240 (2007).
126. Hass, M., Gölnitz, U., Müller, S., Becker-Ziaja, B. & Günther, S. Replicon system for Lassa virus. J Virol 78, 13793–13803 (2004).
127. Bartenschlager, R. & Pietschmann, T. Efficient hepatitis C virus cell culture system: what a difference the host cell makes. Proc Natl Acad Sci USA 102, 9739–9740 (2005).
128. Feigelstock, D. A., Mihalik, K. B., Kaplan, G. & Feinstone, S. M. Increased susceptibility of Huh7 cells to HCV replication does not require mutations in RIG-I. 1–8 (2010). doi:10.1186/1743-422X-7-44
129. Pietschmann, T. et al. Construction and characterization of infectious intragenotypic and intergenotypic hepatitis C virus chimeras. Proc Natl Acad Sci USA 103, 7408–7413 (2006).
130. Wilkins, T., Malcolm, J. K., Raina, D. & Schade, R. R. Hepatitis C: diagnosis and treatment. Am Fam Physician 81, 1351–1357 (2010).
131. Hahn, von, T. et al. Hepatitis C virus continuously escapes from neutralizing antibody and T-cell responses during chronic infection in vivo. Gastroenterology 132, 667–678 (2007).
132. Neumann, A. U. et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science 282, 103–107 (1998).
133. Gale, M. & Foy, E. M. Evasion of intracellular host defence by hepatitis C virus. Nat Cell Biol 436, 939–945 (2005).
134. Szöllösi, J. et al. Fluorescence energy transfer measurements on cell surfaces: a critical comparison of steady-state fluorimetric and flow cytometric methods. Cytometry 5, 210–216 (1984).
135. Horváth, G. et al. Selecting the right fluorophores and flow cytometer for fluorescence resonance energy transfer measurements. Cytometry A 65, 148–157 (2005).
136. Nagy, P. et al. Novel calibration method for flow cytometric fluorescence resonance energy transfer measurements between visible fluorescent proteins. Cytometry A 67, 86–96 (2005).
137. Shapiro, H. Practical Flow Cytometry. 1–47 (2003).
138. Dainiak, M. B., Kumar, A., Galaev, I. Y. & Mattiasson, B. Methods in cell separations. Adv. Biochem. Eng. Biotechnol. 106, 1–18 (2007).

139. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. Annalen der Physik 6, 55–75 (1948).
140. Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. Proc Natl Acad Sci USA 58, 719–726 (1967).
141. Truong, K. & Ikura, M. The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo. Curr Opin Struct Biol 11, 573–578 (2001).
142. Kahra, D. et al. Conformational Plasticity and Dynamics in the Generic Protein Folding Catalyst SlyD Unraveled by Single-Molecule FRET. Journal of Molecular Biology 411, 781–790 (2011).
143. Lorent, E., Bierau, H., Engelborghs, Y., Verheyden, G. & Bosman, F. Structural characterisation of the hepatitis C envelope glycoprotein E1 ectodomain derived from a mammalian and a yeast expression system. Vaccine 26, 399–410 (2008).
144. Michalak, J. P. et al. Characterization of truncated forms of hepatitis C virus glycoproteins. J Gen Virol 78 (Pt 9), 2299–2306 (1997).
145. Lee, J. E., Fusco, M. L. & Ollmann Saphire, E. An efficient platform for screening expression and crystallization of glycoproteins produced in human cells. Nat Protoc 4, 592–604 (2009).
146. Birnboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513–1523 (1979).
147. Mullis, K. et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. 51 Pt 1, 263–273 (1986).
148. Graham, F. L. & Van der Eb, A. J. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52, 456 (1973).
149. Sambrook, J. & Russell, D. W. Calcium-phosphate-mediated Transfection of Eukaryotic Cells with Plasmid DNAs. CSH Protoc 2006, (2006).
150. Kato, T. et al. Cell culture and infection system for hepatitis C virus. Nat Protoc 1, 2334–2339 (2006).
151. Mosmann, T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J. Immunol. Methods 65, 55–63 (1983).
152. Tiselius, A. Electrophoresis of serum globulin. Biochem J 31, 313–317 (1937).
153. Krey, T. et al. The disulfide bonds in glycoprotein E2 of hepatitis C virus reveal the tertiary organization of the molecule. PLoS Pathog 6, e1000762 (2010).
154. Steinmann, E. et al. Hepatitis C Virus p7 Protein Is Crucial for Assembly and Release of Infectious Virions. PLoS Pathog 3, e103 (2007).
155. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27, 431–432 (2011).
156. de Chassey, B. et al. Hepatitis C virus infection protein network. Molecular Systems Biology 4, 230 (2008).
157. Pedamallu, C. S. & Posfai, J. Open source tool for prediction of genome wide protein-protein interaction network based on ortholog information. Source Code Biol Med 5, 8 (2010).
158. Aizaki, H., Lee, K.-J., Sung, V. M. H., Ishiko, H. & Lai, M. M. C. Characterization of the hepatitis C virus RNA replication complex associated with lipid rafts. Virology 324, 450–461 (2004).
159. Love, R. A., Brodsky, O., Hickey, M. J., Wells, P. A. & Cronin, C. N. Crystal structure of a novel dimeric form of NS5A domain I protein from hepatitis C virus. J Virol 83, 4395–4403 (2009).
160. Love, R. A. et al. The conformation of hepatitis C virus NS3 proteinase with and without NS4A: a structural basis for the activation of the enzyme by its cofactor. Clin Diagn Virol 10, 151–156 (1998).
161. Zhu, H. & Briggs, J. M. Mechanistic role of NS4A and substrate in the activation of HCV NS3 protease. Proteins 79, 2428–2443 (2011).
162. Bartenschlager, R. The NS3/4A proteinase of the hepatitis C virus: unravelling structure and function of an unusual enzyme and a prime target for antiviral therapy. J Viral Hepat 6, 165–181 (1999).
163. Op de Beeck, A. et al. The transmembrane domains of hepatitis C virus envelope glycoproteins E1 and E2 play a major role in heterodimerization. J Biol Chem 275, 31428–31437 (2000).
164. Deleersnyder, V. et al. Formation of native hepatitis C virus glycoprotein complexes. J Virol 71, 697–704 (1997).
165. Lorenz, I. C., Allison, S. L., Heinz, F. X. & Helenius, A. Folding and dimerization of tick-borne encephalitis virus envelope proteins prM and E in the endoplasmic reticulum. J Virol 76, 5480–5491 (2002).
166. Wu, S.-R., Haag, L., Sjöberg, M., Garoff, H. & Hammar, L. The dynamic envelope of a fusion class II virus. E3 domain of glycoprotein E2 precursor in Semliki Forest virus provides a unique contact with the fusion protein E1. J Biol Chem 283, 26452–26460 (2008).
167. Steinmann, E. & Pietschmann, T. Hepatitis C virus p7-a viroporin crucial for virus assembly and an emerging target for antiviral therapy. Viruses 2, 2078–2095 (2010).
168. Vieyres, G. et al. Subcellular Localization and Function of an Epitope-Tagged p7 Viroporin in Hepatitis C Virus-Producing Cells. J Virol 87, 1664–1678 (2013).
169. GONZALEZ, M. Viroporins. FEBS Lett 552, 28–34 (2003).
170. Nieva, J. L., Madan, V. & Carrasco, L. Viroporins: structure and biological functions. Nat Rev Microbiol 10, 563–574 (2012).
171. Lai, C. K., Jeng, K. S., Machida, K. & Lai, M. M. C. Association of Hepatitis C Virus Replication Complexes with Microtubules and Actin Filaments Is Dependent on the Interaction of NS3 and NS5A. J Virol 82, 8838–8848 (2008).
172. Hahn, von, T. Virale Infektion von Hepatozyten. Dtsch med Wochenschr 137, 2448–2452 (2012).
173. Da Costa, D. et al. Reconstitution of the Entire Hepatitis C Virus Life Cycle in Nonhepatic Cells. J Virol 86, 11919–11925 (2012).
174. Tai, A. W. et al. A Functional Genomic Screen Identifies Cellular Cofactors of Hepatitis C Virus Replication. Cell Host and Microbe 5, 298–307 (2009).
175. Ng, T. I. et al. Identification of host genes involved in hepatitis C virus replication by small interfering RNA technology. Hepatology 45, 1413–1421 (2007).
176. Randall, G. et al. Cellular cofactors affecting hepatitis C virus infection and replication. Proc Natl Acad Sci USA 104, 12884–12889 (2007).
177. Itsui, Y. et al. Expressional screening of interferon-stimulated genes for antiviral activity against hepatitis C virus replication. J Viral Hepat 13, 690–700 (2006).
178. Supekova, L. et al. Identification of human kinases involved in hepatitis C virus replication by small interference RNA library screening. J Biol Chem 283, 29–36 (2008).
179. Kim, S.-S. et al. A cell-based, high-throughput screen for small molecule regulators of hepatitis C virus replication. Gastroenterology 132, 311–320 (2007).
180. Siegel, R. M. Measurement of Molecular Interactions in Living Cells by Fluorescence Resonance Energy Transfer Between Variants of the Green Fluorescent Protein. Science's STKE 2000, 1pl–1 (2000).
181. Piston, D. W. & Kremers, G.-J. Fluorescent protein FRET: the good, the bad and the ugly. Trends in Biochemical Sciences 32, 407–414 (2007).
182. Chen, H., Puhl, H. L., Koushik, S. V., Vogel, S. S. & Ikeda, S. R. Measurement of FRET Efficiency and Ratio of Donor to Acceptor Concentration in Living Cells. Biophysical Journal 91, L39–L41 (2006).
183. Koushik, S. V., Blank, P. S. & Vogel, S. S. Anomalous surplus energy transfer observed with multiple FRET acceptors. PLoS ONE 4, e8031 (2009).
184. Maeda, T. & Nishimura, S.-I. FRET-Based Direct and Continuous Monitoring of Human Fucosyltransferases Activity: An Efficient synthesis of Versatile GDP-L-Fucose Derivatives from Abundantd-Galactose. Chem. Eur. J. 14, 478–487 (2008).
185. Kim, J., Lee, J., Kwon, D., Lee, H. & Grailhe, R. A comparative analysis of resonance energy transfer methods for Alzheimer related protein–protein interactions in living cells. Mol. BioSyst. 7, 2991 (2011).
186. Koppensteiner, H., Banning, C., Schneider, C., Hohenberg, H.

& Schindler, M. Macrophage Internal HIV-1 Is Protected from Neutralizing Antibodies. J Virol 86, 2826–2836 (2012).
187. Kühl, A. et al. The Ebola virus glycoprotein and HIV-1 Vpu employ different strategies to counteract the antiviral factor tetherin. Journal of Infectious Diseases 204 Suppl 3, S850–S860 (2011).
188. Thyrock, A., Stehling, M., Waschbüsch, D. & Barnekow, A. Characterizing the interaction between the Rab6 GTPase and Mint3 via flow cytometry based FRET analysis. Biochem Biophys Res Commun 396, 679–683 (2010).
189. Somvanshi, R. K., Chaudhari, N., Qiu, X. & Kumar, U. Heterodimerization of β2 adrenergic receptor and somatostatin receptor 5: Implications in modulation of signaling pathway. J Mol Signal 6, 9 (2011).
190. Asbach, B., Ludwig, C., Saksela, K. & Wagner, R. Comprehensive Analysis of Interactions between the Src-Associated Protein in Mitosis of 68 kDa and the Human Src-Homology 3 Proteome. PLoS ONE 7, e38540 (2012).
191. Hassinen, A. et al. Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. Journal of Biological Chemistry 286, 38329–38340 (2011).
192. Olsen, M. J. et al. Function-based isolation of novel enzymes from a large library. Nat. Biotechnol. 18, 1071–1074 (2000).
193. King, F. W., Liszewski, W., Ritner, C. & Bernstein, H. S. High-throughput tracking of pluripotent human embryonic stem cells with dual fluorescence resonance energy transfer molecular beacons. Stem Cells Dev. 20, 475–484 (2011).
194. Selby, M., Glazer, E., Masiarz, F. & Houghton, M. Complex Processing and Protein:Protein Interactions in the E2:NS2 REgion of HCV. Virology 114–122 (1994).
195. Mizushima, H., Hijikata, M., Tanji, Y., Kimura, K. & Shimotohno, K. Analysis of N-terminal processing of hepatitis C virus nonstructural protein 2. J Virol 68, 2731–2734 (1994).
196. Krey, T. Supplementary_Information. 1–1 (2010).
197. Patel, J., Patel, A. H. & McLauchlan, J. The transmembrane domain of the hepatitis C virus E2 glycoprotein is required for correct folding of the E1 glycoprotein and native complex formation. Virology 279, 58–68 (2001).

9 Supplement

Interaction and Prediction Databases:

Databases & Data Collections
Experimental Data

ASEdb; Alanine Scanning Energetics DataBase; db of hotspots in 3D protein structures
Bacteriome.org (University of Toronto); bacterial protein interaction db; integrating physical (protein-protein) and functional interactions within the context of an E. coli knowledgebase
BID Wiki
BIND Biomolecular INteraction Network Database at the University of Toronto, Canada. No species restriction
Binding Interface Database; organize vast amounts of protein interaction information into tabular form,; graphical contact maps, and descriptive functional profiles
BioGRID (Samuel Lunenfeld Research Institute); General Repository for Interaction Datasets; db of genetic and physical interactions
BOND (Thomson Corp.); Biomolecular Object Network Databank; new resource to perform cross-database searches of available sequence, interaction, complex and pathway information; integrates a range of component databases including Genbank and BIND, the Biomolecular Interaction Network Database
Campylobacter jejuni Interactions Database (Wayne State University); includes protein interaction data from a large-scale yeast two-hybrid (YTH) screen and interactions predicted from experimental data in other organisms (interologs)
CYGD PPI section of the Comprehensive Yeast Genome Database. Manually curated comprehensive S. cerevisiae PPI database at MIPS
DIP (UCLA) Database of Interacting Proteins at UCLA. No species restriction.
DOMINO - domain peptide interactions database, describing interactions mediated by protein-interaction domains
DroID (Wayne State University) DROsophila Interactions Database; comprehensive gene and protein interactions database designed specifically for the model organism Drosophila
EchoBASE (University of York); integrated post-genomic database for E. coli
GRID General Repository for Interaction Datasets. Mount Sinai Hospital, Toronto, Canada
GWIDD (University of Kansas); Genome WIde Docking Database; combines available experimental data with models built by docking techniques contains known protein-protein interactions and allows input of other sequences and structures to find interacting proteins and obtain the structure of their complexes
HCPIN - Human Cancer Pathway Protein Interaction Network (Rutgers University); constructed by analysis of several classical cancer-associated signaling pathways and their; physical protein-protein interactions
HIV-1 - Human Protein Interaction Database (NCBI); summary of all known interactions of HIV-1 proteins with host cell proteins, other HIV-1 proteins, or proteins from disease organisms associated with HIV / AIDS
hp-DPI (National Health Research Institutes); Helicobacter Pylori Database of Protein Interactomes; combined with experimental and inferring interactions
HPID (Inha University); Human Protein Intercation Database; Department of computer Science and Information Engineering Inha University, Inchon, Korea
HPRD The Human Protein Reference Database. Institute of Bioinformatics, Bangalore, India and Johns Hopkins University, Baltimore, MD, USA.
HUGE ppi (Kazusa DNA Research Institute) Human Unidentified Gene-Encoded large proteins; db of protein-protein interactions between large KIAAproteins
Human Protein Reference Database (Johns Hopkins University & The Institute of Bioinformatics, India); platform to visually depict and integrate information pertaining to domain architecture, post-translational modifications, interaction networks and disease association for each protein in the human proteome
I2D (Ontario Cancer Institute); Interologous Interaction Database of known and predicted mammalian and eukaryotic protein-protein interactions
IBIS (NCBI) Inferred Biomolecular Interactions Server; reports physical interactions observed in experimentally-determined structures for a given protein; infers/predicts interacting partners and binding sites by homology
ICBS (University of California) Inter-Chain Beta-Sheets database; protein-protein interactions mediated by interchain ß-sheet formation
KDBI (National University of Singapore); db of Kinetic Data of Bio-molecular Interactions
KEGG BRITE (Kyoto University) Biomolecular Relations in Information Transmission and Expression functional hierarchies and binary relationships of biological entities
MetaCore Commercial software suite and database. Manually curated human PPIs (among other things). GeneGo
MINT (Centro di Bioinformatica Moleculare, Universita di Roma, Italy) Molecular INTeractions database; db of functional interactions between biological molecules: RNA, DNA, proteins
molmovdb.org (Yale University) db of macromolecular movements with associated tools for flexibility and geometric analysis
MPact (MIPS); yeast protein-protein interaction data contained in the Comprehensive Yeast Genome Database (CYGD)
MPIDB (J. Craig Venter Institute); Microbial Protein Interaction DataBase; provide all known physical microbial interactions; experimentally determined interactions among proteins of 250 bacterial species/strains
MPPI (MIPS); Mammalian Protein-Protein Interaction database; collection of manually curated high-quality PPI data collected from the scientific literature by expert curators
MRC PPI links Commented list of links to PPI databases and resources maintained at the MRC Rosalind Franklin Cetre for Genomics Research, Cambridge, UK
NetPro (Molecular Connections); database of protein-protein and protein-small molecules interaction consisting of more than 320,000 interactions captured from more than 1500 abstracts, approximately 1600 published journals and more than 60,000 references
OPHID The Online Predicted Human Interaction Database. Ontario Cancer Institute and University of Toronto, Canada
PathCalling Proteomics and PPI tool/database. CuraGen Corporation.
Pawson Lab Information on protein-interaction domains.
PDZBase (Weill Medical College of Cornell University); manually curated protein-protein interaction database developed specifically for interactions involving PDZ domains; currently contains 339 experimentally determined protein-protein interactions
PepCyber: (University of Minnesota); database of human protein-protein interactions mediated by phosphoprotein binding domains (PPBDs)
PINT (bioinfodatabase.com); Protein-protein INteraction Thermodynamic db; contains experimental data of several thermodynamic parameters along with literature, sequence and structural information and experimental conditions
POINT (National Health Research Institutes & National Taiwan University); functional database for the prediction of the human protein-protein interactome based on available orthologous interactome datasets integrates several publicly accessible databases, with emphasis placed on the extraction of a large quantity of mouse, fruit fly, worm and yeast protein-protein interactions datasets from the Database of Interacting Proteins (DIP), followed by conversion of them into a predicted human interactome
PRIME (Human Genome Center, University of Tokyo); PRotein Interaction and Molecular Information database; integrated gene/protein informatics database based on natural language processing
Protein Interaction Maps - PIMs (Hybrigenics); functional proteomics software platform, dedicated to the exploration of protein pathways; PIM's available for Helicobacter pylori, Hiv_Human, Drosophila and TGF-Beta Hybrigenics PPI data and tool, H. pylori. Free academic license available
Protein-Protein Interaction Panel using mouse full-length cDNAs (RIKEN, Yokohama Institute); see Suzuki et al., Genome Res. 2001, 11, 1758-1765
Protein-Protein Interaction Server Analysis of protein-protein interfaces of protein complexes from PDB. University College of London, UK.
ProtoArray® (Invitrogen)

PSIbase (BioSystems Dept., KAIST & BiO centre); molecular interaction database focuses on structural interaction of proteins and their domains
PutidaNET (Korean BioInformation Center); proteome database of Pseudomonas putida KT2440; provides predicted protein-protein interaction, gene ontology information, and physio-chemical information
RIKEN Experimental and literature PPIs in mouse.
SNAPPIView (University of Dundee); Structures, iNterfaces and Alignments for Protein-Protein Interactions; object-oriented database of domain-domain interactions observed in structural data
SPiD (INRA) db of two-hybrid protein interactions in B. Subtilis
SPIN-PP Server (Columbia University); Surface Properties of INterfaces - Protein Protein interfaces database of all protein-protein interfaces in the PDB
VirusMint - Virus protein interactions db
Yeast Interacting Proteins Database (Kanazawa University); yeast protein interactome; view data with a Genetic Network Visualization System
Yeast Protein Linkage Map Data (University of Washington)
YPD™ (BIOBASE) Yeast Proteome Database; comprehensive knowledge resource for the proteins of S. cerevisiae "BioKnowledge Library" at Incyte Corporation. Manually curated PPI data from S. cerevisiae. Proprietary.

Databases & Data Collections
Predictions

ADAN *(EMBL);* prediction of protein-protein inter**A**ction of mo**D**ular dom**A**i**N**s
***Arabidopsis* Interactions Viewer** *(CSB University of Toronto);* db of predicted and confirmed *Arabidopsis* interacting proteins
AtPID *(Northeast Forest University); **A**rabidopsis **T**haliana* **P**rotein **I**nteractome **D**atabase*;* centralized platform to depict and integrate the information pertaining to protein-protein interaction networks, domain architecture, ortholog information and GO annotation in the Arabidopsis thaliana proteome*;* integrates data from several bioinformatics prediction methods and manually collected information from the literature
cons-PPISP *(Florida State University);* consensus protein-protein interaction site predictor consensus neural network method for predicting protein-protein interaction sites
Fly-DPI *(National Health Research Institutes); Drosophila melanogaster* database of protein interactomes statistical model to predict protein interaction networks
Genes2Networks *(Mount Sinai School of Medicine);* connecting lists of gene symbols using mammalian protein interactions databases powerful web-based software that can help to interpret lists of genes and proteins can be used to find relationships between genes and proteins from seed lists, and predict additional genes or proteins that may play key roles in common pathways or protein complexes
HAPPI *(Indiana University School of Informatics, Purdue University School of Science);* **H**uman **A**nnotated and **P**redicted **P**rotein **I**nteraction db collected or inferred computationally from public sources
INTERPARE *(National Genome Information Center, Korea Research Institute of Bioscience and Biotechnology & BiO Centre)* protein interfaceome database contains large-scale interface data of proteins with known 3D-structures
meta-PPISP *(Florida State University);* meta web server for protein-protein interaction site prediction
JCB PPI site at the Jena Centre for Bioinformatics, Germany
MitoInteractome *(Korean Bioinformation Center);* web-based portal containing information on predicted protein-protein interactions, physicochemical properties, polymorphism, and diseases related to the mitochondrial proteome contains 6,549 protein sequences which were extracted from the following databases: SwissProt, MitoP, MitoProteome, HPRD and Gene Ontology database
NOXclass *(Max-Planck-Institut für Informatik);* SVM (support vector machine algorithm) classifier identifying protein-protein interaction types
PIBASE *(University of California);* comprehensive database of structurally defined interfaces between pairs of protein domains
PIPs *(University of Dundee);* human protein-protein interaction prediction
PPIDB *(Iowa State University);* database of protein-protein interfaces derived from all protein-protein complexes available in the Protein Data Bank
Predictome *(Boston University);* database of putative links between proteins using sequence data of genomes of 71 microorganisms*;* Predicted functional associations and interactions. Boston University.
PRIMOS *(BIOMIS, FH Hagenberg);* **PR**otein **I**nteraction and **MO**lecule **S**earch db; knowledge portal for analysing protein-protein interaction data
PRISM *(Koc University);* **PR**otein **I**nteractions by **S**tructural **M**atching; explore protein interfaces and predict protein-protein interactions
PRODISTIN Web Site *(LGPD/IBDM, CNRS);* web service to functionally classify genes/proteins from any type of interaction network
Prolinks Database *(University of California);* collection of inference methods used to predict functional linkages between proteins methods include the *Phylogenetic Profile* method, which uses the presence and absence of proteins across multiple genomes to detect functional linkages; the *Gene Cluster* method, which uses genome proximity to predict functional linkage; *Rosetta Stone*, which uses a gene fusion event in a second organism to infer functional relatedness; and the *Gene Neighbor* method, which uses both gene proximity and phylogenetic distribution to infer linkage
Protein Interaction Network of *E. coli* *(Centre for DNA Fingerprinting and Diagnostics);* obtained by training a Support Vector Machine on the high quality of interactions in the EcoCyc database, and with the assumption that the periplasmic and cytoplasmic proteins may not interact with each other
SNAPPI-Predict *(University of Dundee);* **S**tructures, i**N**terfaces and **A**lignments for **P**rotein-**P**rotein **I**nteractions protein-protein interaction prediction program
SPIDer *(Beijing Normal University); **S**accharomyces* **P**rotein-protein **I**nteraction **D**atabase effective method of reconstructing a yeast protein interaction network by measuring relative specificity similarity (RSS) between two Gene Ontology (GO) terms
SynechoNET *(Korean BioInformation Center);* integrated protein-protein interaction database of a model cyanobacterium *Synechocystis* sp. PCC 6803 shows feasible cyanobacterial domain-domain interactions, as well as their protein level interactions; provides transmembrane topology and domain information, as well as the interaction networks in graphical web interfaces
Databases & Data Collections
Related Domain, Pathway and Network Databases
BioCarta (*BioCarta*) charting pathways of life
BioCyc Database Collection *(SRI);* collection of 506 pathway/genome databases each database describes the genome and metabolic pathways of a single organism
BN++ *(Center for Bioinformatics Saar & Center for Bioinformatics Tübingen);* biochemical network library for analyzing and visualizing complex biochemical networks and processes
CellCircuits *(University of California);* open-access db of molecular network models that bridges the gap between databases of individual pair-wise molecular interactions and databases of validated pathways contains functional network hypothesis produced by algorithms that screen molecular interaction networks based on their correspondence with expression or phenotypic data, their internal structure, or their conservation across species
DAPID *(National Chiao Tung University);* **D**omain **A**nnotated **P**rotein-protein **I**nteraction **D**atabase; db of domain-annotated protein interactions inferred from three-dimensional (3D) interacting domains of protein complexes in the Protein Data Bank (PDB)
DIMA *(MIPS, TUM);* **D**omain **I**nteraction **Ma**p*;* comprehensive resource for functional and physical interactions among conserved protein-domains
DOMINE *(University of Texas at Dallas)* db of known and predicted protein domain (domain-domain) interactions
DOQCS *(NCBS)* **D**atabase **O**f **Q**uantitative **C**ellular **S**ignaling
EDGEdb *(University of Massachusetts Medical School) C. **E**legans* **D**ifferential **G**ene **E**xpression **d**ata**b**ase
EMP *(EMP Project Inc.)* **E**nzymes and **M**etabolic **P**athways
HotSprint *(KOC University, Turkey)* db of computational hot spots in protein interfaces
iHOP *(Computational Biology Center, Memorial Sloan-Kettering Cancer Center, USA & Protein Design Group, National Center of Biotechnology, Spain);* **I**nformation **H**yperlinked **O**ver **P**roteins; Protein association network built by literature mining
InterDom *(Laboratories for Information Technology;* Institute for InfoComm Research, Singapore.*);* db of putative **INTER**acting protein **DOM**ains derived from multiple sources
KEGG; Kyoto **E**ncyclopedia of **G**enes and **G**enomes

KEGG LIGAND; database of chemical compounds and reactions in biological pathways

Kinase Pathway Database *(Human Genome Center);* integrated database concerning completed sequenced major eukaryotes, which contains the classification of protein kinases and their functional conservation and orthologous tables among species, protein-protein interaction data, domain information, structural information, and automatic pathway graph image interface

Negatome Database *(MIPS);* collection of protein and domain pairs which are unlikely engaged in direct physical interactions contains experimentally supported non-interacting protein pairs derived from two distinct sources: by manual curation of literature and by analysing protein complexes from the PDB can be used to evaluate newly derived experimental interactions

PATHWAY Database *(ProteinLounge);* largest collection of signal transduction and metabolic pathways including extensive reviews and detailed protein information

Pfam *(Sanger Institute);* **P**rotein **FAM**ilies database of alignments and HMM

PPISearch *(National Chiao Tung University);* web server for searching homologous protein-protein interactions across multiple species

SCOPPI *(TU Dresden);* **S**tructural **C**lassification **O**f **P**rotein-Protein **I**nterfaces*;* db of all domain-domain interactions and their interfaces derived from PDB structure files and SCOP domain definitions

SCOWLP *(BIOTEC TU Dresden);* **S**tructural **C**haracterization **O**f **W**ater, **L**igands and **P**roteins web application represent a framework for the study of protein interfaces and comparative analysis of protein family binding regions

SMART *(EMBL Heidelberg);* **S**imple **M**odular **A**rchitecture **R**esearch **T**ool

SPAD *(Kyushu University);* **S**ignaling **PA**thway **D**atabase*;* integrated db for genetic information and signal transduction systems

TRANSCompel *(BIOBASE)* db of composite regulatory elements affecting gene transciption in eukaryotes

TRANSPATH *(BIOBASE);* db on molecular pathways and cellular network modeling

UniHI *(Charite - Medical Devision, Humboldt-University zu Berlin);* **Uni**fied **H**uman **I**nteractome*;* comprehensive database of the computational and experimental based human protein interaction networks

The Interactive Fly *(Society for Developmental Biology);* a guide to *Drosophila* genes and their role in development includes information on **biochemical pathways**

Wnt Signaling Pathway *(Stanford University Medical Center);* resource for members of the Wnt community, providing information on progress in the field, maps on signaling pathways, methods and various other items

Yeast Pathways in the *Comprehensive Yeast Genome Database (MIPS)*

APID *(Cancer Research Center* Salamanca, Spain*);* **A**gile **P**rotein **I**nteraction **D**ataAnalyzer interactive web-tool; all known experimentally validated protein-protein interactions

eFsite *(Osaka University);* **E**lectrostatic surface of **F**unctional-**SITE***;* db for molecular surfaces of proteins functional sites, displaying the electrostatic potentials and hydrophobic properties together on the Connolly surfaces of the active sites

Expression Profiler *(European Bioinformatics Institute);* explores protein interaction data using expression data

FunSimMat *(Max Planck Institute for Informatics);* **Fun**ctional **Sim**ilarity **Mat**rix; comprehensive resource of semantic and functional similarity values

IntAct Project *(EMBL-EBI):* protein interaction db and toolkit for the storage, presentation and analysis of protein interactions; no species restriction

InterPreTS *(EMBL);* **INTER**action **PRE**diction through **T**ertiary **S**tructure

InterProSurf *(University of Texas Medical Branch);* web server for predicting the functional sites on protein surfaces

InterViewer *(Inha University);* visualization of large-scale protein interaction networks

iSPOT *(Universita di Roma)* prediction of protein-protein interactions mediated by families of peptide recognition modules

Medusa *(EMBL);* interface to the **STRING** protein interaction db*;* a general graph visualization tool

PathBLAST *(Whitehead Institute);* network alignment and search tool for comparing protein interaction networks across species to identify protein pathways and complexes that have been conserved by evolution

PEDANT *(GSF);* **P**rotein **E**xtraction, **D**escription, and **An**alysis **T**ool*;* exhaustive automatic analysis of genomic sequences by a large variety of bioinformatics tools

PDBSiteScan *(Institute of Cytology and Genetics SBRAS);* designed for searching 3D protein fragments similiar in structure to known active, binding and posttranslational modification sites

PIMRiderTM *(Hybrigenics);* proteomic software and interaction data

PIMWalkerTM *(Hybrigenics);* a free protein-protein interaction map java viewer

PIVOT *(Tel Aviv University)* **P**rotein **I**nteractions **V**isualizati**O**n **T**ool

ProFace *(Department of Biochemistry, Bose Institute);* server for the analysis of the physicochemical features of protein-protein interfaces suite of programs that uses a file, containing atomic coordinates of a multi-chain molecule, as input and analyzes the interface between any two or more subunits

Protein3D Home *(LECB)*

PROTORP *(University of Sussex);* protein-protein interface analysis server analyse the properties of interfaces in the 3D structures of protein-protein associations

SCOWLP *(TU Dresden);* **S**tructural **C**haracterization **O**f **W**ater, **L**igands and **P**roteins*;* web-based relational db describing PDB interface interactions at atom, residue and domain level

SPIN-PP Server *(Columbia University);* **S**urface **P**roperties of **IN**terfaces - **P**rotein **P**rotein interfaces*;*

STRING *(EMBL);* **S**earch **T**ool for the **R**etrieval of **IN**teracting **G**enes/proteins*;* database of known and predicted protein-protein interactions for a large number of organisms interactions include direct (physical) and indirect (functional) associations data are derived from four sources: genomic context, high-throughput experiments, (conserved) coexpression and previous knowledge

YETI *(University Edinburgh);* **Y**east **E**xploration **T**ool **I**ntegrator*;* workbench tool for visualization/analysis of post-genomic data sets available for S. *cerevisiae*

Danksagung

Mein außerordentlicher Dank geht an Dr. Michael Schindler für die Ermöglichung und die Betreuung der vorliegenden Arbeit. Ebenso möchte ich mich besonders bei meinen Betreuern Prof. Dr. Thomas Dobner und PD Dr. Markus Perbandt bedanken, welche mir immer mit Rat und Tat zu Seite standen sowie die Erstellung der Gutachten und die Durchführung der Disputation übernommen haben. Für das Übernehmen des Vorsitzes dieser danke ich Prof. Dr. Alexander Haas.

Der ehemaligen Nachwuchsgruppe für Viruspathogenese danke ich für die wirklich schöne Zeit im gemeinsamen Labor. Hierbei geht ein besonderer Dank an Carina, Kristin und Herwig die mich auf meinem gesamten Weg der Doktorarbeit begleiten durften, sowie an unsere ‚Neuen' Karen und Stephan, welche mir alle Freunde fürs Leben geworden sind.

Der gesamten Abteilung für Molekulare Virologie, danke ich für die freundliche Aufnahme meinerseits in ihrer Gruppe um mir ein Beenden der Doktorarbeit – im Rahmen der Graduiertenschule SDI – zu ermöglichen. Im Besonderen danke ich hierbei Melanie Schmid für das Korrekturlesen vorliegender Arbeit. Mein Dank gilt ebenso dem gesamten Heinrich-Pette-Institut und seinen Mitarbeitern für die Hilfsbereitschaft, die mir in den letzten 3 ½ Jahren entgegengebracht wurde.

Dem gesamten SDI-Team möchte ich für die wirklich schöne Zeit im Rahmen einer neu gewonnenen Familie danken, welche mir viele Einblicke in vollkommen neue Themengebiete aber auch in die unterschiedlichsten Persönlichkeiten gegeben hat.

Des Weiteren danke ich Dr. Marta Kotasinska (UKE) für ihre stete Hilfsbereitschaft in sämtlichen Belangen der Proteinreinigung als auch Dr. Frank Lennartz von der Philipps Universität Marburg (Institute of Virology) für die Expertise bezüglich dem Drosophila-Expressions-System.

Vor allem danke ich Bernd Schimmer für die erholsamen Feierabende, für den seelisch-moralischen Ausgleich vom Arbeits-Alltag sowie für seine Unterstützung und Hilfe in jeder Hinsicht während der gesamten Studienzeit.
Meiner Familie Kerstin, Tatjana und Gerda danke ich für das Ermöglichen meines Studiums, nicht nur aus finanzieller Sicht. Sie haben mich, in dem was ich tue, immer bestätigt und an mich geglaubt.