Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2007) 546–557
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Structural and functional comparison of the non-structural protein 4B in flaviviridae

Christoph Welsch [a,b,*], Mario Albrecht [b], Jochen Maydt [b], Eva Herrmann [a], Martin Walter Welker [a], Christoph Sarrazin [a], Axel Scheidig [c], Thomas Lengauer [b], Stefan Zeuzem [a]

[a] *Internal Medicine II, Saarland University Hospital, Kirrberger Straße, 66421 Homburg/Saar, Germany*
[b] *Max Planck Institute for Informatics, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany*
[c] *Structural Biology, Saarland University, Kirrberger Straße, 66421 Homburg/Saar, Germany*

Received 6 October 2006; received in revised form 23 March 2007; accepted 28 March 2007
Available online 4 April 2007

Abstract

Flaviviridae are evolutionarily related viruses, comprising the hepatitis C virus (HCV), with the non-structural protein 4B (NS4B) as one of the least characterized proteins. NS4B is located in the endoplasmic reticulum membrane and is assumed to be a multifunctional protein. However, detailed structure information is missing. The hydrophobic nature of NS4B is a major difficulty for many experimental techniques. We applied bioinformatics methods to analyse structural and functional properties of NS4B in different viruses. We distinguish a central non-globular membrane portion with four to five transmembrane regions from an N- and C-terminal part with non-transmembrane helical elements. We demonstrate high similarity in sequence and structure for the C-terminal part within the flaviviridae family. A palmitoylation site contained in the C-terminal part of HCV is equally conserved in GB virus B. Furthermore, we identify and characterize an N-terminal basic leucine zipper (bZIP) motif in HCV, which is suggestive of a functionally important interaction site. In addition, we model the interaction of the bZIP region with the recently identified interaction partner CREB-RP/ATF6β, a human activating transcription factor involved in ER-stress. In conclusion, the versatile structure, together with functional sites and motifs, possibly enables NS4B to adopt a role as protein hub in the membranous web interaction network of virus and host proteins. Important structural and functional properties of NS4B are predicted with implications for ER-stress response, altered gene expression and replication efficacy.

Keywords: *In silico* analysis; Flaviviridae; NS4B; Leucine zipper; Structure model

1. Introduction

The flaviviridae family comprises the genera flaviviruses, pestiviruses, hepaciviruses and unclassified flaviviridae, the GB-virus group. The genus flavivirus is an arthropod-borne virus, which can cause severe diseases like yellow-fever or dengue-fever and dengue hemorrhagic fever in humans. These diseases are of high medical relevance, especially in moist savanna zones of West and Central Africa and in parts of South America (yellow-fever), as well as in tropical and subtropical regions in Southeast Asia, South and Central America and several Carribean Islands (dengue-fever, http://www.cdc.gov/ncidod/dvbid/). In contrast, only tick-borne encephalitis is common in Europe and in North America. Pestiviruses, such as the bovine-viral diarrhea virus, play a role in veterinary medicine.

Hepaciviruses are a growing health-care problem with about 200 million people infected with hepatitis C virus worldwide. Chronic hepatitis C can lead to cirrhosis and its sequelae, including the development of hepatocellular carcinoma [1–3]. The hepatitis C virus genome is about 9.6 kB in length and encodes a polyprotein of at least four structural (core, envelope proteins E1 and E2, p7) and six non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B), cleaved by cellular as well as viral proteases [4–6].

* Corresponding author at: Saarland University Hospital, Internal Medicine II, Kirrberger Straße, Building 41, 66421 Homburg/Saar, Germany.
Tel.: +49 6841 16 23581; fax: +49 6841 16 23583.
E-mail addresses: christoph.welsch@uniklinikum-saarland.de, christophwelsch@gmx.net (C. Welsch).

1093-3263/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2007.03.012

Flaviviridae are single positive-strand RNA viruses. The genetic organization as well as general physical properties, such as hydrophobic profiles of the viral polyprotein, are comparable [7,8]. Only little is known about the non-structural protein 4B (NS4B) in flaviviridae [9]. In hepatitis C virus, NS4B comprises four to five transmembrane regions (TMRs) and is located within the endoplasmic reticulum (ER) membrane [10,11]. The protein is thought to be part of a multiprotein complex, which was discovered to be involved in membranous web formation [12]. Using co-immunoprecipitation, NS4B has been shown to be associated with the non-structural proteins NS3 and NS4A in hepatitis C, when expressed in eukaryotic cells. Here, NS4A is essential for interaction with the NS3 protease and the polyprotein substrate NS4B-5A [13]. In bovine viral diarrhea virus NS4B could be chemically cross-linked to the non-structural proteins NS3 and NS5A [14]. The NS5A protein of the hepatitis C virus is modified during the virus life cycle by phosphorylation. Hyperphosphorylation of NS5A in hepatitis C occurs in the cytoplasm and depends on NS4B, amongst other HCV proteins [15]. Several observations concerning the role of NS4B in signal transduction cascades were obtained in cell culture for hepatitis C virus, e.g. activation of NF-κB associated signals, AP-1 and serum responsive element SRE [16,17]. Recently, a nucleotide-binding motif, which binds and hydrolyzes GTP, was described in NS4B of hepaciviruses. GTPase activity seems to be of importance for viral RNA replication [18]. Due to the overall hydrophobic nature of NS4B, detailed structure determination using experimental techniques, e.g. X-ray or NMR spectroscopy, is a sophisticated venture and probably not successful in the foreseeable future. An *in silico* approach seemed to be reasonable for gaining insight into structure and function of NS4B while circumventing hard-to-implement experimental procedures.

2. Methodology

Sequences for flaviviridae NS4B proteins were retrieved from the Pfam database [19] and additionally searched for in the NR Database provided by NCBI [20] and in the UniProt database [21] using the BLAST suite of programs [22]. The NCBI taxonomy browser has been employed for distinguishing different taxa of the flaviviridae family (http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi). We included hepatitis C virus JA (HCV type JA), a well-described HCV genotype 1b prototype strain [23] into our analysis and used this sequence as a reference in Section 3, together with NS4B sequences from GB-viruses A, B and C (GBV type A, B, C), yellow fever virus (YFV), dengue virus type 1, 2 and 4 (DV type 1, 2, 4), Japanese encephalitis virus (JEV), Kunjin virus (KV), tick-borne encephalitis virus (TBEV), classical swine fever virus (CSFV) and bovine viral diarrhea virus (BVDV). The UniProt accession numbers of the sequences reported in this paper are given in Table 1.

2.1. Sequence analysis, transmembrane and secondary structure prediction

Sequence alignments were computed using CLUSTAL W [24] and MUSCLE [25], subsequently improved by minor manual modifications using the SEAVIEW alignment editor [26]. The multiple sequence alignment was prepared using

Table 1
Flaviviridae NS4B sequences

Entry name	UniProt	Origin	Taxonomy	Reference
POLG_HCVJA	P26662	Hepatitis C virus (isolate Japanese)	Hepacivirus	[23]
O39927_9HEPC	O39927	Hepatitis C virus (type 6a)	Hepacivirus	[64]
POLG_HCV1	P26664	Hepatitis C virus (isolate 1)	Hepacivirus	[8]
POLG_HCVH	P27958	Hepatitis C virus (isolate H)	Hepacivirus	[65]
POLG_HCVBK	P26663	Hepatitis C virus (isolate BK)	Hepacivirus	[66]
POLG_HCVJT	Q00269	Hepatitis C virus (isolate HC-JT)	Hepacivirus	[67]
POLG_HCVTW	P29846	Hepatitis C virus (isolate Taiwan)	Hepacivirus	[68]
POLG_HCVJ6	P26660	Hepatitis C virus (isolate HC-J6)	Hepacivirus	[69]
POLG_HCVJ8	P26661	Hepatitis C virus (isolate HC-J8)	Hepacivirus	[70]
O41892_9FLAV	O41892	Hepatitis GB virus A; GBV-A	Unclassified flaviviridae	[71]
Q96898_9FLAV	Q96898	Hepatitis GB virus A; GBV-A	Unclassified flaviviridae	[72]
Q9QEW5_GBVB	Q9QEW5	Hepatitis GB virus B; GBV-B	Unclassified flaviviridae	[73]
Q9DXH4_9FLAV	Q9DXH4	Hepatitis G virus (isolate PEI); GBV-C/HGV group	Unclassified flaviviridae	[74]
POLG_DEN4	P09866	Dengue virus (type 4); dengue virus group	Flavivirus	[75]
POLG_KUNJM	P14335	Kunjin virus (strain MRM61C); Japanese encephalitis virus group	Flavivirus	[76]
POLG_TBEVH	Q01299	Tick-borne encephalitis (strain Hypr); tick-borne encephalitis virus group	Flavivirus	[77]
Q9YRV3_9FLAV	Q9YRV3	Yellow fever virus; yellow fever virus group	Flavivirus	[78]
Q8JQE3_9FLAV	Q8JQE3	Dengue virus (type 1); dengue virus group	Flavivirus	[79]
Q9YKL3_9FLAV	Q9YKL3	Dengue virus (type 2); dengue virus group	Flavivirus	[80]
Q9WCX7_9FLAV	Q9WCX7	Japanese encephalitis virus; Japanese encephalitis virus group	Flavivirus	[81]
POLG_BVDVC	Q96662	Bovine viral diarrhea virus (strain CP7) (mucosal disease virus)	Pestivirus	[82]
POLG_CSFVA	P19712	Classical swine fever virus (strain Alfort) (Hog cholera virus)	Pestivirus	[83]

Different NS4B sequences from flaviviridae (hepaciviruses, GB-viruses, flavi- and pestiviruses) extracted from the UniProt database by a sequence search using the NCBI taxonomy browser. The table contains the sequence entry name, UniProt accession number, origin of the sequence, taxonomy and references for the nucleotide sequences.

GeneDoc [27]. NS4B secondary-structure prediction was performed using three state-of-the-art methods, SamT99 [28], SSPro [29] and PsiPred [30]. Majority voting was applied to the set of secondary structure predictions to compute a reliable overall consensus prediction [31]. Similarly, we computed consensus predictions of the transmembrane regions (TMRs) using the programs TMHMM2 [32], MEMSAT [33] and HMMTOP2 [34]. We used DisEMBL v1.5 (http://dis.embl.de/) [35] to analyse NS4B for intrinsic protein disorder and DILIMOT (http://dilimot.embl.de/) [36] to identify linear motifs.

Hydropathy plots were calculated using the Kyte–Doolittle method with a window size of 19 residue positions (http://fasta.bioch.virginia.edu/fasta_www/grease.htm). A hydropathy histogram was calculated using an own MATLAB 6 routine (http://www.mathworks.com/). Amino acids were classified according to the Black-and-Mould hydrophobicity scale (hydrophilic: D, E, H, K, N, Q, R; neutral: G, S, T; hydrophobic: A, C, F, I, L, M, P, V, W, Y). The EMBL coiled-coil prediction server (http://www.russell.embl-heidelberg.de/cgi-bin/coils-svr.pl) was applied to analyse NS4B for possible coiled-coils using the treewin mode and MTK matrix. The algorithm analyses residues for their probability to participate in coiled-coil structures by comparing the flanking sequences with sequences of known coiled-coil proteins [37]. Coiled-coils are strongly amphipathic and display a pattern of residues with hydrophilic and hydrophobic properties, which repeats every seven residues. A basic leucine zipper (bZIP) is formed by α-helices in coiled-coil formation, with an amino acid heptad repeat, labelled ‘a’ to ‘g’, a hydrophobic core and leucine residues at the ‘d’ positions [37]. We used WebLogo (http://weblogo.berkeley.edu/logo.cgi) for computing a bZIP motif sequence logo for HCV [38].

2.2. Structure prediction

To model the 3D protein structure of the bZIP interaction of CREB-RP/ATF6β with HCV NS4B, we used pairwise sequence-structure alignments with helical elements of the template structure 2ZTA, a synthetic polypeptide corresponding to the leucine zipper of the yeast transcriptional activator GCN4. Each alignment has been submitted to the WHAT IF (http://swift.cmbi.kun.nl/WIWWWI/) homology modeling server. The protein structure images were drawn in the Accelrys Discovery Studio ViewerLite (http://www.accelrys.com/).

2.3. Phylogenetic analysis

We applied ProtTest [39–41] to find the best-fitting model of protein evolution for the flaviviridae sequence alignment. The RtRev model [42] with gamma rates and variable amino-acid frequencies was the best-fitting model according to both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) [43]. We used PhyML [41] for estimating a maximum-likelihood phylogeny for the given protein evolution model and performed 100 bootstrap replicates in order to obtain bootstrap support values.

3. Results

The maximum-likelihood phylogenetic analysis of the flaviviridae family revealed four distinct clusters: hepatitis C virus subtypes with differences in the NS4B gene sequence, GB-viruses, the different flavivirus groups and the pestiviruses, BVDV and CSFV (Fig. 1). HCV is shown as an evolutionary intermediate between GB-viruses and the remainder, with hepatitis C virus subtypes much closer to GB-viruses than to the other taxa. The cluster of dengue, Japanese encephalitis, yellow fever and tick-borne encephalitis viruses is well resolved, as indicated by the respective bootstrap values. The other clusters, particularly hepatitis C virus subtypes, are far less well resolved. Also, the clusters of hepatitis C and GB-viruses cannot be separated with high bootstrap support. Especially the GB-virus B is closely related to hepatitis C. Considering the distances induced by the phylogeny, the GB-viruses are closely related to the cluster of flaviviruses and HCV. The low bootstrap support of 30 also shows that it is hard to discern both clades phylogenetically. The phylogenetic tree emanates a clear difference of pestiviruses from all other investigated viruses, hepaci-, GB- and other flaviviruses. Because of rather low similarity, we excluded pestiviruses from further detailed analysis.

The multiple sequence alignment reveals minor differences within different subtypes of the hepacivirus group. Flaviviruses represent the most heterogeneous group of sequences and can be clearly distinguished from sequences of hepaciviruses (Fig. 2). The protein architecture of NS4B consists of several cytosolic and transmembrane α-helices with a tripartite organization: the N- and C-terminal ends are in the cytoplasm, while the central region is inserted into the ER membrane.

The N-terminal part in the multiple sequence alignment is rather diverse in different flavivirus taxa, whereas the GB-virus sequences exhibit similarities to hepaciviruses. Nevertheless, the N-terminal part is the most diverse region of the NS4B protein in the flaviviridae sequence alignment. The membrane-associated region of NS4B exhibits only minor differences between viral species. Only the N-terminal part of the membrane-associated region shows significant differences between hepaci- and GB-viruses on the one hand and flaviviruses on the other hand. Similarities within the whole flaviviridae family exist in the C-terminal part (Fig. 2).

3.1. N-terminal part of NS4B

The NS4B N-terminal part of hepaciviruses consists of 73 residues (Fig. 2). A physical interaction of the N-terminal part of HCV NS4B (aa 1–62) and the central portion of the human protein CREB-RP/ATF6β (aa 350–388), an ER-stress response element, has been reported previously [44].

Our method predicts helices but no TMRs for all investigated virus strains in the N-terminal part. Results are summarized in an annotated multiple sequence alignment and a domain architecture picture (Figs. 2 and 4B).

In hepaciviruses (e.g. HCV-JA), an α-helix of 61 residues length (aa 5–65) is indicated by a consensus of secondary

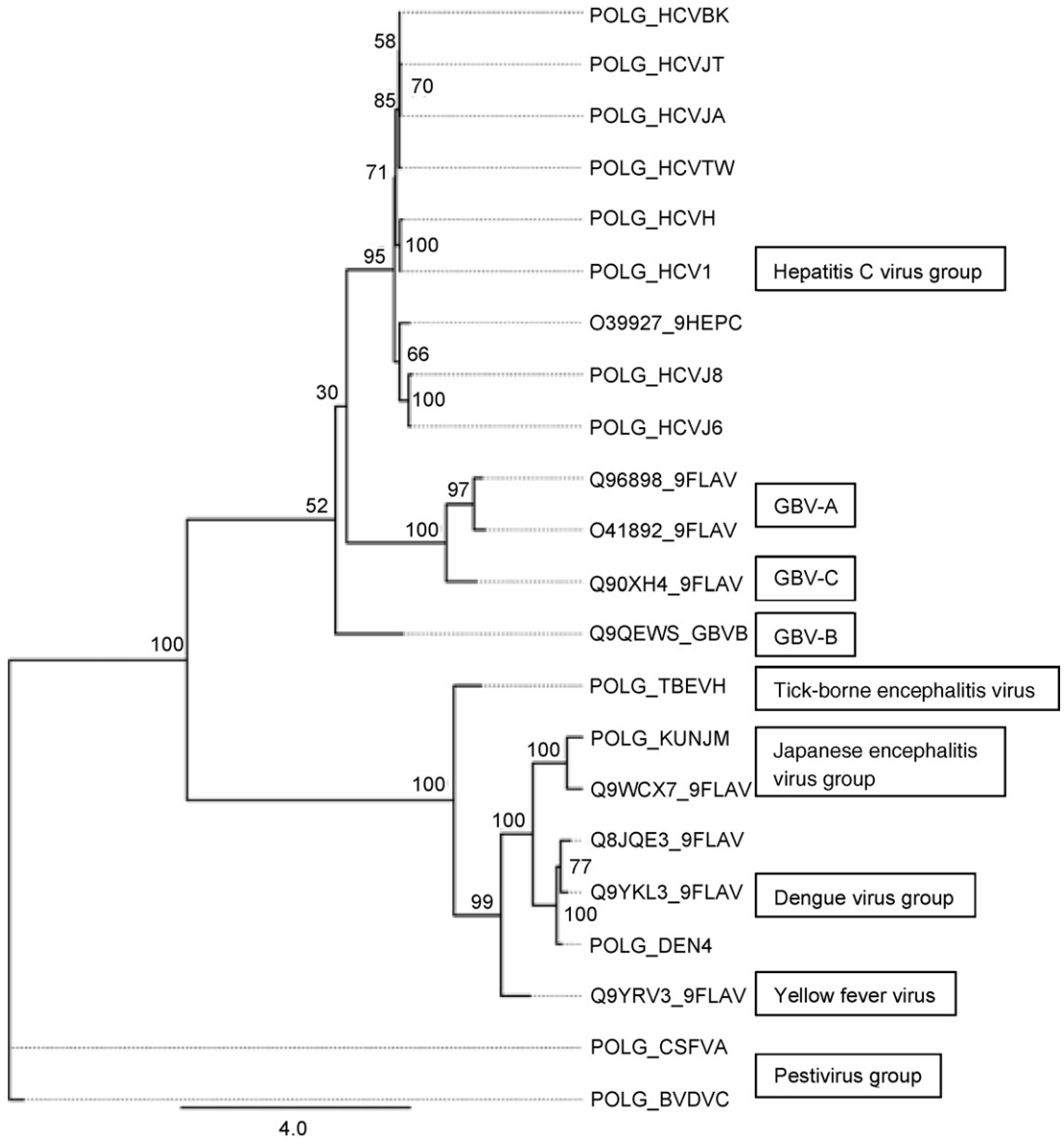

Fig. 1. A maximum-likelihood phylogeny showing the evolutionary relationships of the non-structural protein 4B within different taxa of the flaviviridae family (Table 1): hepaciviruses, GB-viruses, flavi- and pestiviruses. Bootstrap values for interpretation of the phylogenetic tree are given at the internal nodes. High bootstrap values indicate best data support and consistency for a taxon bipartition at that point.

structure prediction methods. Hepacivirus subtypes are well conserved. All of them show a similar pattern of physicochemical properties, representing an amphipathic helix (Fig. 3). An analogous helical element with clear amphipathic properties is missing in any of the other flaviviruses, which show one short α-helix of 12 residues length. The GB-virus N-terminal part is related to hepaciviruses. Nevertheless, an α-helix with amphipathic properties has not been found in the N-terminal region. Instead, two separate α-helices with 28 and 31 residues, respectively, have been predicted.

A hydropathy plot for all investigated HCV subtypes, including residues 22–49, according to the prototype sequence of HCV-JA, reveals amphipathic physicochemical properties and a tandem heptad repeat of residues characteristic of a basic leucine zipper (bZIP) motif [45]. The zipper heptad repeat labelled ‘a’ to ‘g’, usually possesses apolar amino acids at positions ‘a’ and ‘d’, and aliphatic amino acids such as alanine, phenylalanine, isoleucine, leucine, valine or the aromatic amino acid tryptophane at the remaining positions [45]. The suspected viral bZIP motif contains only two of four possible leucines at position ‘d’ (Figs. 2, 5 and 6) in HCV-JA. On the other hand, seven of eight heptad positions in ‘a’ or ‘d’ are occupied by apolar amino acids aiding in the formation of a well-conserved hydrophobic core (Figs. 5A and 6). The formation of heterodimeric coiled-coils by bZIP motifs of two different proteins is often stabilized by ionic interactions (‘‘salt bridges’’). This is due to charged amino acids (histidine, lysine, arginine, aspartate, glutamate) at positions ‘e’ and ‘g’, bordering the leucine zipper hydrophobic core [46–49]. We found in HCV-JA that the bZIP motif contains one lysine at

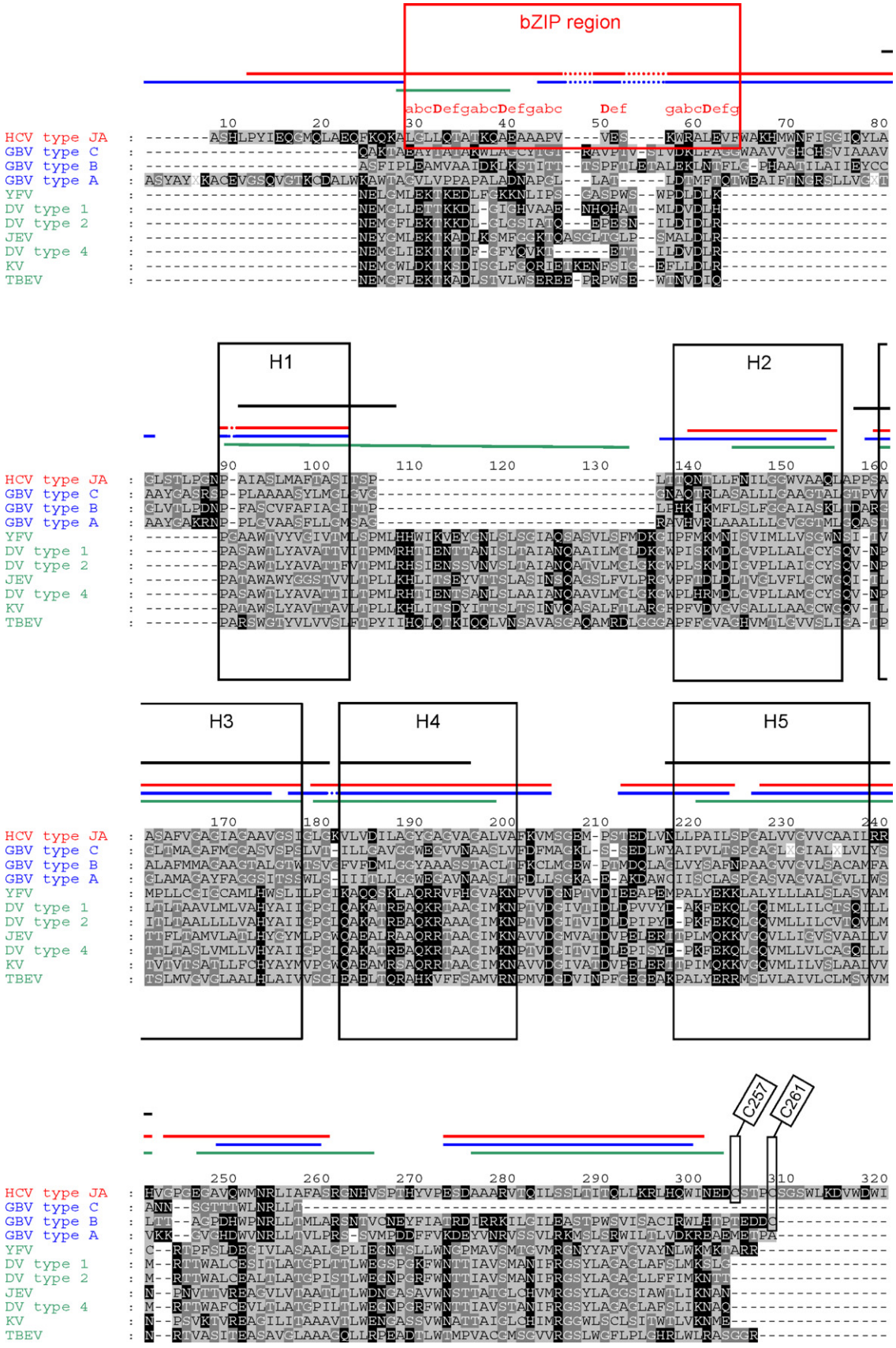

Fig. 2. Multiple sequence alignment of the non-structural protein 4B (NS4B) for members of the flaviviridae family, hepatitis C virus prototype sequence JA, GB-viruses A, B and C and different flaviviruses. Sequences were taken from the UniProt database and referenced by their accession numbers. Abbreviations used are the following (see Section 2): HCV type JA, GB types A, B and C, YFV, DV types 1, 2 and 4, JEV, KV and TBEV. Amino-acid residues are shaded in different grey levels,

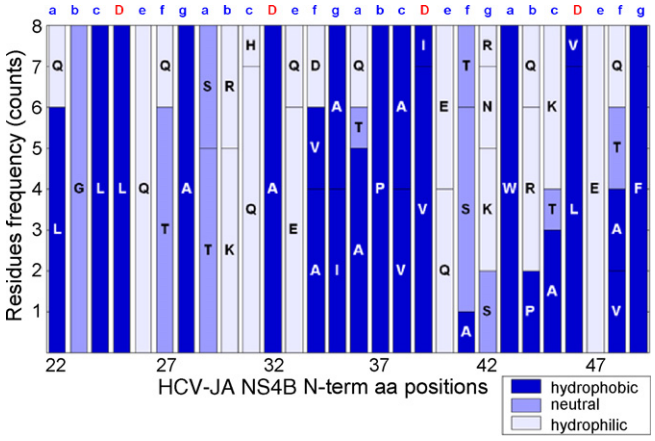

Fig. 3. Hydropathy histogram of the suspected bZIP motif in eight different HCV subtypes (hepaciviruses, Table 1), coloured according to hydrophobicity (Black-and-Mould scale). Each amino acid as well as the corresponding frequency is shown.

position 'e' and three glutamates at position 'g', corresponding to four of eight heptad positions containing amino acids potentially involved in ionic interactions (Figs. 2 and 5A). The N-terminal part (residues 0–50) of HCV-JA NS4B showed a higher coil probability than the protein rest. No coiled-coil was predicted for the membrane-associated part of NS4B, and only a very low probability for the C-terminal part of NS4B.

In addition, we computed a sequence logo based on a multiple alignment of 229 hepatitis C virus sequences, taken from the Pfam database. The sequence logo has been used to investigate for conservation at crucial amino-acid positions of the suspected bZIP motif. The region investigated is relatively diverse. However, positions accounting for the bZIP motif showed high conservation. The allocation of amino-acid physicochemical properties seems to be preserved (Fig. 5B). A corresponding bZIP motif cannot be identified in GB-viruses and flaviviruses (Fig. 2). For the hepatitis C virus, an interaction of the NS4B basic leucine zipper with the cytoplasmic domain of the eukaryotic protein CREB-RP/ATF6β, an activating transcription factor, has been observed experimentally by Tong et al. [44]. Using the bZIP motif, we modeled this interaction *in silico* by helical wheels [50] (Fig. 5A). They demonstrated conserved hydrophobic residues building the hydrophobic core, two leucine–leucine interactions (Leu25 and Leu46 from HCV NS4B and equivalent residues from CREB-RP/ATF6β) and one ionic-interaction by Glu47 from HCV NS4B and Arg377 from CREB-RP/ATF6β. Residues 28–38 within the suspected bZIP motif of HCV-JA are predicted of low-complexity by the intrinsic protein disorder predictor DisEMBL. A motif search using DILIMOT did not detect any linear motifs in the N-terminal part of NS4B as well as in the rest of the protein.

3.2. Membrane-associated part of NS4B

The central part of the flaviviridae NS4B is thought to be responsible for integration of the protein into the endoplasmic reticulum (ER) membrane. Similar to the N-terminal part of NS4B, the membrane-associated part of hepaciviruses is highly conserved and the GB-viruses are more closely related to hepaciviruses than flaviviruses. In comparison to the N-terminal part, the sequence alignment seems to be more conserved (Fig. 2).

Secondary structure prediction proposes four to five helices numbered H1 to H5 in the membrane-associated part of the NS4B protein. For hepaciviruses and GB-viruses, all of them (H1, H2, H3, H4 and H5, Fig. 4) have been predicted as transmembrane regions (TMRs) by hydrophobicity plots and TM-prediction programs. H1, in comparison to H2, H3, H4 and H5, shows by far the lowest value in terms of hydrophobicity (Kyte–Doolittle plot, Fig. 4A) and is therefore the most speculative of the TMRs in the current literature [11]. The consensus prediction for the first α-helix (H1) comprises residues 74–86. A TMR has been predicted for hepaci- (HCV-JA), GB- and flaviviruses as well. The TM-prediction for HCV-JA includes residues 63–86, thus containing a region (aa 66–73), which has been predicted of low-complexity by secondary structure prediction programs. As a conclusion of results of secondary structure and TM-predictions, as well as hydropathy plots within all flaviviridae, the TMR-consensus for the first α-helix H1 excludes the low-complexity region. This region represents a flexible loop in the structure of NS4B, connecting an amphipathic helix containing the bZIP motif of the N-terminal globular part of NS4B and helix H1 of the membrane-associated part. The second α-helix (H2) contains residues 92–109. The prediction is the same for HCV-JA and the GB-viruses, whereas, in the group of flaviviruses, the α-helix is predicted consistently to be five positions shorter at the N-terminal part. A TMR for the second α-helix (H2) is only predicted for GB-virus A. Nevertheless, the hydrophobic profile for that region (Fig. 4A) suggests a transmembrane orientation of H2. A glycosylation assay by Lundin et al. [11] has been performed to investigate NS4B with respect of the location and orientation of TMRs. Within this study, a transmembrane orientation of H2 has been postulated. Five residues separate H1 and H2. The third α-helix (H3) ranges from residue 114 to residue 131 and is concordantly predicted as a TMR. A fourth α-helix (H4) prediction includes residues 136–154. Helix prediction is in agreement for all investigated

according to their physicochemical properties (Black-and-Mould scale). The alignment is numbered at the top of each row. Secondary structure predictions of α-helices, according to majority voting results, are indicated by coloured lines (HCV, red; GB-viruses, blue; flaviviruses, green). A consensus by majority voting for flaviviridae TM-predictions is indicated by a black line. Putative transmembrane helices are indicated by black boxes, referenced H1 to H5. Beginning and ending positions of a box reflect a trade-off by taking into account secondary structure predictions, TM-predictions, hydropathy plots and considering residues that favour (E, A, L and M) or disfavour (P, G, C and T) the presence of an α-helix. The bZIP motif is framed by a red box and amino acids are referenced 'a' to 'g', according to the heptad repeats. D in upper case indicates the leucine position within a regular zipper motif. Palmitoylation sites at Cys257 and Cys261 (according to HCV-JA) are indicated in the C-terminal part.

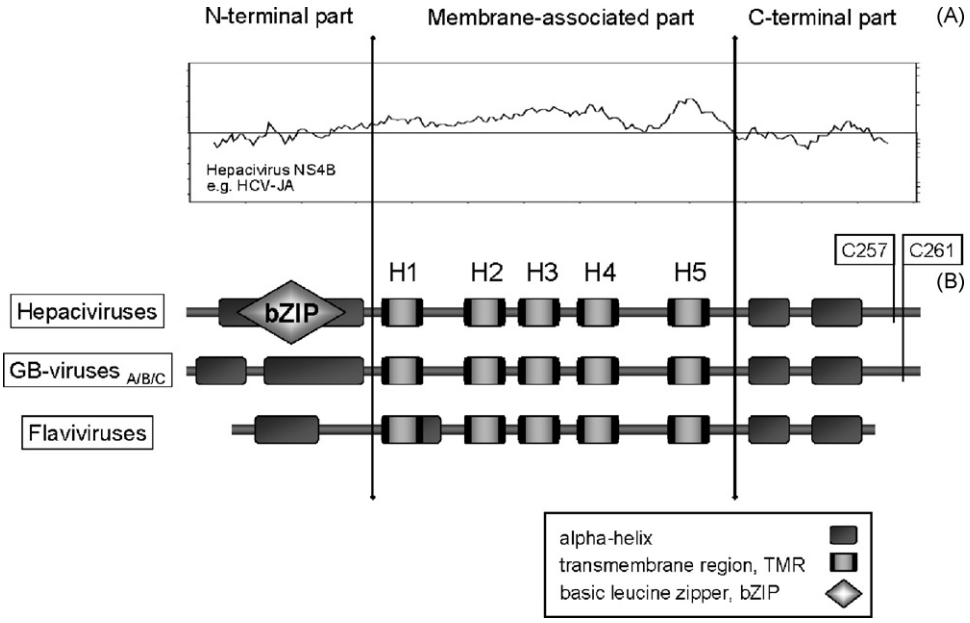

Fig. 4. (A) Kyte–Doolittle hydrophobicity plot of HCV-JA NS4B using a window size of 19 residues and (B) domain architectures of NS4B homologs of the flaviviridae family (hepatitis C viruses, GB-viruses and other flaviviruses, e.g. tick-borne encephalitis virus, TBEV). Palmitoylation sites (Cys257 and Cys261) and a suspected site in HCV-JA and GB-virus B, respectively, are indicated in the C-terminal part.

viruses, whereas TM-predictions differ significantly in length and position. In flaviviruses, H4 has not been predicted as a TMR. H2, H3 and H4 are equally spaced from each other by short segments of low-complexity, which represent helix connecting loops, of four residues each. The fifth helix (H5) is consistently predicted as a TMR and runs from residue 171 to residue 191. It shows the highest hydrophobicity values in comparison among all TMRs in the membrane-associated part. A sequence segment from residues 155–170, including several gaps, is separating H4 and H5. In comparison to the other segments of four to five residues separating H1, H2, H3 and H4, this segment is a rather large, ER-lumen oriented loop comprising 16 residues.

A recent publication described the existence of a nucleotide-binding motif and GTPase activity in HCV NS4B [18]. Investigating the flaviviridae multiple sequence alignment (Fig. 2) and a Pfam sequence alignment of all available hepatitis C virus NS4B proteins, including 476 sequences (alignment not shown), for the previously reported conserved nucleotide-binding motifs, Walker A and Walker B [18], no amino-acid pattern is striking [51,52].

The Walker A motif, the most conserved element of a nucleotide-binding domain, typically looks as follows: GxxxxGK(T/S). Here x stands for any amino acid. Glycine, lysine and threonine and serine, respectively, are highly conserved and of functional importance. In our study, no such sequence motif has been observed, neither in hepaciviruses, nor in any of the GB- or flaviviruses investigated. A Pfam alignment (not shown) of 238 hepatitis C virus sequences, including the suspected region for the Walker A motif, revealed that 7.6% of the lysines were substituted by arginine and all threonines and serines were replaced by the chemically dissimilar amino acids valine (89%) and isoleucine (11%), respectively. The calculated consensus sequence of the proposed Walker A motif is 'xxxxxGxx' (100% consensus), with only glycine conserved. For example, lysine is replaced by arginine in the hepatitis C virus type 6a, UniProt accession number O39927. The three final residues of the Walker A motif are the minimum of what is required for proper function. In particular, the lysine is needed to interact with the phosphate groups of MgATP/MgADP, and threonine or serine at the C-terminal end of the Walker A motif are involved in the ligation of the Mg^{2+} ion [53]. The Walker B motif ideally consist of four residues, 'D(P/E)xx'. The aspartic acid is important for coordinating the Mg^{2+} ion in MgATP/MgADP [53]. In NS4B of hepatitis C virus subtypes investigated (Table 1), no such sequence motif was found to be conserved. In particular, Walker A and B motifs are characteristic of the NTPase fold, which mainly consists of a central β-sheet [54]. However, no β-strand is predicted in NS4B of flaviviridae.

3.3. C-terminal part of NS4B

Our results show two α-helices but no TMRs in the C-terminal protein portion (Figs. 2 and 4B). The first α-helix is 13 residues in length (aa 200–213) and the second α-helix comprises 23 residues (aa 229–252). Both secondary structure elements are interrupted by an intermediate low-complexity region with 15 residues. Neither the first nor the second α-helix has been predicted as a TMR. Throughout flaviviridae, the C-terminal part is similar, especially with respect to the predicted secondary structure elements. A yet unknown common function for that part of the protein in the flaviviridae family can be assumed (Figs. 2 and 4B).

Recently, lipid modification important for protein polymerization has been reported for the C-terminal part of NS4B.

Table 2
HCV-JA transmembrane helix predictions

TMHMM2	64–86	101–123	136–158	168–190
MEMSAT	63–86	114–133	140–158	168–191
HMMTOP2	69–92	105–128	141–160	173–196

The table lists the sequence positions returned by three different transmembrane helix prediction servers (TMHMM2, MEMSAT, HMMTOP2) for the complete NS4B protein of HCV-JA.

Palmitoylation was demonstrated for hepatitis C virus NS4B on cysteine residues 257 and 261, according to HCV-JA [23]. Site-directed mutagenesis revealed the importance of the palmitoylated residues, particularly Cys261, for protein–protein interaction [55]. The multiple sequence alignment (Fig. 2) revealed cysteine residues in HCV-JA at the very C-terminal end, next to the cytosolic helical elements. We could not find any transmembranous element for HCV-JA neighboring the palmitoylation sites. Our transmembrane predictions are summarized in Table 2. Regarding the flaviviridae family, cysteine residues in the C-terminal part of NS4B have been found in GB-virus B (three times), tick-borne virus, dengue types 1, 2 and 4 (one time), Japanese encephalitis viruses (one to two times), but not in yellow-fever virus, GB-virus A or GB-virus C. Interestingly only the GB-virus B, which has been identified by phylogeny as the most similar virus to the hepatitis C virus clade, revealed a cysteine residue conserved in the Cys261 equivalent position. Therefore, NS4B driven lipid modification and polymerization can be discussed for GB-virus B, similar to HCV.

4. Discussion

We used an *in silico* approach to predicting the structure and function of the non-structural protein 4B (NS4B) by comparing different members of the flaviviridae family. A phylogenetic tree of representative viruses revealed four different clades. Hepaciviruses, GB-viruses, flaviviruses and pestiviruses can be distinguished as expected. The fact that GB-viruses are more closely related to hepaciviruses than flaviviruses is noteworthy. Hepatitis C virus subtypes (hepaciviruses) and GB-viruses cannot be separated with high bootstrap support for assessing statistical significance. Especially the GB-virus B could not be separated clearly from hepaciviruses and might also cluster together with hepatitis C viruses. Likewise, the sequence alignment and structure prediction show more agreement between hepaciviruses and GB-viruses than between hepaciviruses and flaviviruses. The pestiviruses BVDV and CSFV clustered together, well differentiated from all other investigated viruses, represented in the phylogenetic tree. They showed very low similarity in NS4B to all other viruses investigated.

The N-terminal part of NS4B seems to be the most diverse. A previously described amphipathic α-helix [56] has been identified for hepaciviruses. By means of co-immunoprecipitation and yeast two-hybrid assays, an interaction of the N-terminal part of NS4B with the bZIP motif of the eukaryotic ER-stress response element CREB-RP/ATF6β has been demonstrated previously [44].

We identified a tandem heptad repeat of 28 residues exclusively in hepatitis C virus, located in the central part of the amphipathic α-helix, with typical physicochemical properties corresponding to a basic leucine zipper motif (bZIP) (Fig. 2). The suspected motif is not a typical bZIP domain, because only two of four possible leucines at position ‘d’ (Leu25 and Leu46) are present and conserved in HCV-JA. In agreement with this finding, the EMBL coiled-coil prediction server predicted a higher coil probability for the N-terminal part of HCV NS4B. Nevertheless, the fact that a fully automated procedure could not identify the virus bZIP motif may due to the fact that most prediction methods efficiently work for bacterial or eukaryotic proteins. In contrast, viral protein sequences are more flexible in evolutionary terms and thus conserved motifs are harder to detect. Nevertheless, seven of eight residues at positions ‘a’ and ‘d’ are apolar, aliphatic or aromatic and therefore able to form the important hydrophobic core of the leucine zipper. Ionic interactions are often stabilizing, especially in the formation of heterodimeric leucine zippers. Since acidic and basic amino acids are important in ionic interactions, the fact that four of eight residues of the HCV-JA NS4B fulfil this property supports our assumption about the presence of a bZIP domain within the N-terminal part of HCV NS4B (Figs. 2 and 5A).

A Pfam alignment of 229 hepatitis C virus sequences containing the suspected bZIP motif region showed a particularly high degree of conservation of crucial zipper positions (Fig. 5B). Because of low sequence conservation, a similar bZIP domain in GB- or flaviviruses could not be identified. One possible interaction partner for the HCV NS4B leucine zipper has been reported previously [44]. We build a 3D homology-model of the hypothesised leucine–zipper interaction of HCV-JA NS4B with CREB-RP/ATF6β, showing a highly conserved hydrophobic core, which is a major determinant of dimer stability and two of four possible leucine residues (Leu25 and Leu46) facing equivalent residues in CREB-RP/ATF6β (Fig. 6A and B). The only ‘‘salt-bridge’’ observed for that interaction was build by Glu47 from HCV-JA NS4B and Arg377 from CREB-RP/ATF6β. Nevertheless, the fact that four residues at positions ‘e’ or ‘g’ are positively or negatively charged, could enable HCV NS4B to efficiently interact with leucine zipper domains in many other proteins as well. This is of special interest regarding the observation that NS4B is a crucial component of HCV membranous web formation and viral replication. The bZIP motif could be of importance to allow protein–protein interactions with host factors and viral proteins. Since the overall confidence of the 3D model is only moderate, it should be evaluated further using experimental techniques of protein interaction analysis.

CREB-RP/ATF6β is an ER-membrane bound transcription factor, activated upon ER-stress. In response to the accumulation of unfolded or misfolded proteins in the ER-lumen, CREB-RP/ATF6β is cleaved by intramembrane proteolysis to release its cytoplasmic domain. The released domain enters the nucleus in order to regulate gene transcription [57]. The hydrophobic core with leucine residues facing each other and an ionic interaction, which border the hydrophobic core, are major

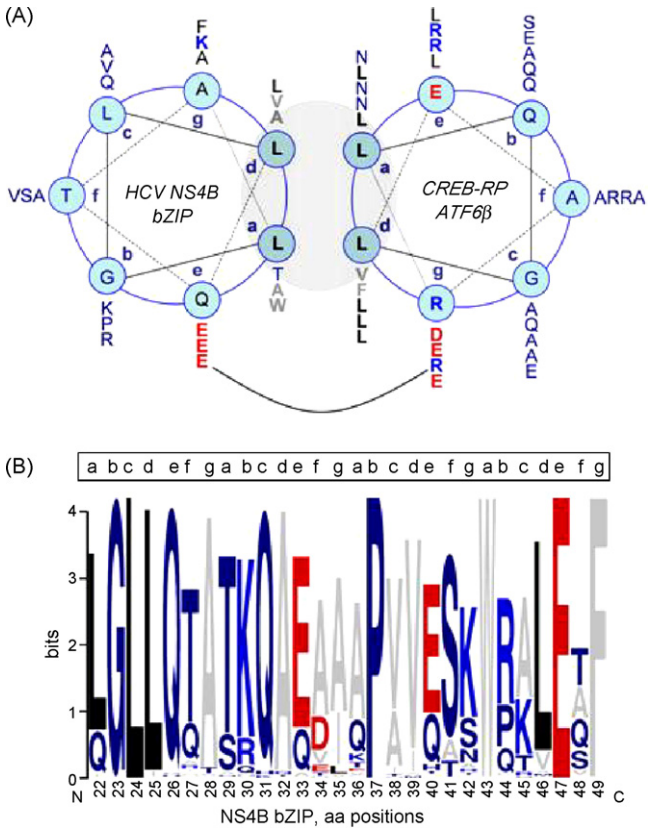

Fig. 5. (A) Proposed *in silico* helical wheel model for the bZIP interaction of a hepatitis C virus NS4B and CREB-RP/ATF6β. Schematic view from the top of the α-helix containing heptad repeats. Grey-shaded hydrophobic core and residues available for ionic interactions in red and light blue, negative (D and E) and positive (H, K and R) charged amino acids, respectively. Amino acids contributing to either the hydrophobic core (A, F, I, V and W) or leucine residues are coloured grey and black, respectively. An ionic interaction in case of the HCV-JA NS4B and CREB-RP/ATF6β leucine zipper is highlighted by a black-arc. (B) Sequence logo derived from 229 hepatitis C virus NS4B sequences contained in the Pfam database showing the suspected bZIP motif region. Amino acids are coloured as follows: A, F, I, V and W in grey (apolar), L in black, D and E in red (negative charged), H, K and R in light blue (positive charged) and C, G, M, N, P, Q, S, T and Y in dark blue (rest).

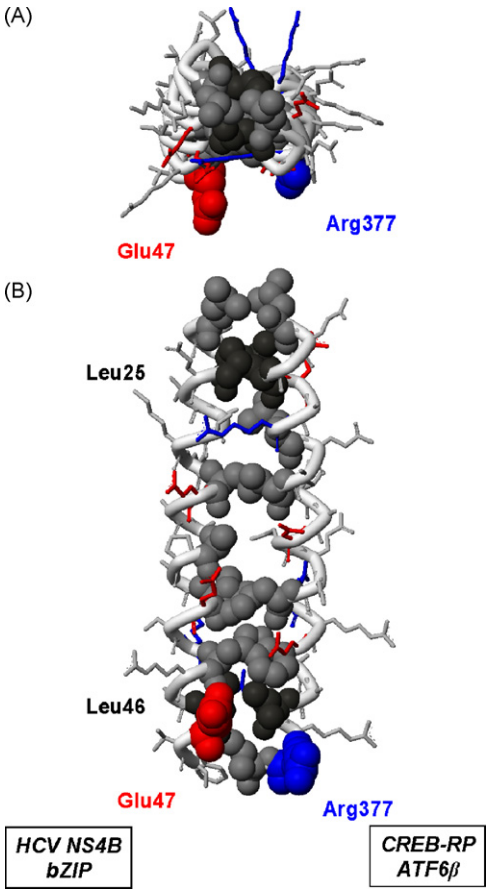

Fig. 6. (A) Top view of the bZIP homology model showing amino acids of the hydrophobic core in CPK mode, coloured grey and black, with leucine–leucine interactions (Leu25 and Leu46, HCV-JA NS4B) shown in black. The ionic interaction of Glu47 (HCV-JA NS4B) and Arg377 (CREB-RP/ATF6β) is shown in CPK mode, red and blue, respectively, according to their negative and positive charge. (B) Corresponding side view of the bZIP 3D model.

determinants of heterodimer stability of the zipper coiled-coil interaction [46–49]. As a consequence, NS4B could link to the cytoplasmic CREB-RP/ATF6β domain and impair the translocation into the nucleus. In this way, the hepatitis C virus could interfere and modulate ER-stress induced gene expression patterns [58–62].

An amino-acid substitution Gln1737His in a hepatitis C virus replicon cell line (Gln26 according to the prototype sequence HCV-JA) has been reported to affect the function of NS4B and to contribute to the development of interferon (IFN) resistance [63]. Interestingly, the reported substitution is located within the suspected bZIP motif described in this work. The substitution alters the fifth heptad position ‘e’, which, according to our model, is potentially involved in ionic interactions (Figs. 2, 5A and 6). The exchange of a neutral glutamine with the basic amino acid histidine enables the formation of an additional ‘‘salt-bridge’’ at a crucial position of the bZIP domain, possibly leading to a more effective heterodimer formation with potential interaction partners. This could explain the observed phenotypic IFN resistance in the replicon cell line and conversely support the presence of the bZIP domain within HCV NS4B (Fig. 2).

For the central part of NS4B, we could confirm four to five TMRs. So far, the first TMR (H1, Fig. 4B) is the most speculative in our study as well as in the current literature [11]. As the protein has to be cleaved by cytoplasmic proteases [4], an initial orientation of the N-terminal part towards the cytoplasm is needed, which supports an NS4B model comprising four TMRs.

By introducing glycosylation sites into the protein, Lundin et al. demonstrated a translocation of the N-terminal tail into the ER-lumen [11]. Two conformational states of the N-terminal TMR are reasonable. Only in case of four TMRs, an N-terminal orientation towards the cytoplasm occurs, which would allow bZIP interaction with the cytoplasmic domain of CREB-RP/ATF6β to take place.

It has been recently shown that NS4B promotes polymerization, maybe important for membranous web formation and viral RNA replication [55]. The polymerization domain has been mapped mainly to the N-terminal 70 amino acids, which

underlines the hypothesised multifunctionality of the basic leucine zipper identified in this report.

A nucleotide-binding motif, which binds and hydrolyzes guanosine triphosphate (GTPase) has been published for hepaciviruses [18]. The authors mentioned high conservation of Walker A and Walker B motifs in different hepatitis C virus isolates. In search of amino-acid consensus patterns in the multiple sequence alignment including GB-viruses and flaviviruses (Fig. 2), as well as in a Pfam full sequence alignment for HCV NS4B, no conserved sequence pattern has been found in the present study. We could show that the lysine in Walker A, which is indispensable for the interaction of the enzyme with phosphate groups of MgATP/MgADP, is only conserved in 92.4% of all sequences. Threonine or serine in Walker A, responsible for ligation of the Mg^{2+} ion, is never conserved in hepatitis C viruses. NTPases share a common structure, with a core of a central pleated, at least four-stranded, parallel β-sheet, flanked by α-helices. The overall shape of the nucleotide-binding site is a characteristic feature in NTPase families and only minor topological differences exist [54]. Besides the lack of conserved amino-acid motifs, we did not predict any β-strand for NS4B. The previously suspected Walker A and B motifs have been described to be located between the two TMRs H3 and H4 of the membrane-associated part of NS4B, and between the two cytoplasmic helices of the globular C-terminal part of NS4B. In particular, the Walker A motif would be in a rather rigid conformation, which does not agree with the need of flexibility in NTPases. Therefore, NS4B GTPase activity seems very unlikely to occur.

The C-terminal part of NS4B revealed significant similarities throughout all investigated viruses of the flaviviridae family. We predicted two α-helical elements, but no TMRs. This part of NS4B is oriented towards the cytoplasm. In our analysis, the consistency of secondary structure elements and sequence similarities points to a yet unknown common function of the C-terminal globular part in flaviviridae.

Lipid modification by palmitoylation of conserved cysteine residues has been demonstrated for several HCV genotypes (Cys257 and Cys261, according to HCV-JA) [55]. Palmitoylation has been discussed as an important factor for NS4B protein–protein interactions in the formation of the HCV RNA replication complex. Although the main part in polymerization resides in the N-terminal part of NS4B, the C-terminal part has been shown to facilitate polymerization [55]. All three different servers consistently predicted transmembrane helices in HCV NS4B with respect to the length of the helices and their sequence positions. However, the large distance between the last transmembrane element and the palmitoylation sites is surprising. Nevertheless, none of the different transmembrane prediction servers found additional transmembrane helices in HCV closer to the reported palmitoylation sites. Interestingly the GB-virus B only, which has been identified as a close HCV homolog, comes with a cysteine in a Cys261 equivalent position. Probably NS4B in GB-virus B, comparable with what has been observed for HCV NS4B, induces specialized membrane structures and could be involved in the RNA replication complex.

5. Conclusions

In conclusion, we suggest that the structure of NS4B has modular domain architecture, possessing transmembrane elements as well as flexible loops and helices. This architecture enables many different orientations of important sites and amino-acid motifs (e.g. the bZIP motif, reported in this work) to enable several protein interactions to occur. The summarized multifunctionality of NS4B can be related to the versatile structure of the protein. Potentially this allows NS4B to become an essential protein hub in the membranous web interaction network of virus and host proteins. The bZIP motif that we detected in the N-terminal part of NS4B calls for further experimental validation. The interacting domain of NS4B with CREB-RP/ATF6β has to be characterized in more detail. Multiple site-directed mutagenesis experiments of the bZIP motif should be conducted to further investigate and characterize effects on the strength of the protein–protein interaction. If an interaction has been clearly confirmed, a luciferase assay should be employed for semi-quantitative analysis of the replication efficacy in the HCV replicon model.

In the future, a complex model of protein interactions and time-dependent gene regulation patterns is needed to better understand intertwined cellular processes during viral infection.

Acknowledgments

This study was supported by a DFG grant to C.W., M.A., E.H., C.S., A.S., T.L. and S.Z. (Klinische Forschergruppe, KFO 129/1-1, TP2, TP3, TP6, TP7) and two European Networks of Excellence funded by the European Commission, viRgil (LSHM-CT-2004-503359) and BioSapiens (LSHG-CT-2003-503265).

References

[1] H.B. El-Serag, Hepatocellular carcinoma: recent trends in the United States, Gastroenterology 127 (2004) S27–S34.
[2] J.H. Hoofnagle, Course and outcome of hepatitis C, Hepatology 36 (2002) S21–S29.
[3] L.B. Seeff, The natural history of chronic hepatitis C virus infection, Clin. Liver Dis. 1 (1997) 587–602.
[4] M. Hijikata, H. Mizushima, Y. Tanji, Y. Komoda, Y. Hirowatari, T. Akagi, N. Kato, K. Kimura, K. Shimotohno, Proteolytic processing and membrane association of putative non-structural proteins of hepatitis C virus, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 10773–10777.
[5] K.E. Reed, C.M. Rice, Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties, Curr. Top. Microbiol. Immunol. 242 (2000) 55–84.
[6] K. Shimotohno, Y. Tanji, Y. Hirowatari, Y. Komoda, N. Kato, M. Hijikata, Processing of the hepatitis C virus precursor protein, J. Hepatol. 22 (1995) 87–92.
[7] Q.L. Choo, J. Han, A.J. Weiner, L.R. Overby, D.W. Bradley, G. Kuo, M. Houghton, Hepatitis C virus is a distant relative of the flaviviruses and pestiviruses, in: T. Shikata, R.H. Purcell, T. Uchida (Eds.), Viral Hepatitis C, D and E, Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1991, pp. 47–52.
[8] Q.L. Choo, K.H. Richman, J.H. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, R. Medina-Selby, P.J. Barr, et al., Genetic organization and diversity of the hepatitis C virus, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 2451–2455.

[9] S. Miller, S. Sparacio, R. Bartenschlager, Subcellular localization and membrane topology of the dengue virus type 2 non-structural protein 4B, J. Biol. Chem. 281 (2006) 8854–8863.
[10] T. Hugle, F. Fehrmann, E. Bieck, M. Kohara, H.G. Krausslich, C.M. Rice, H.E. Blum, D. Moradpour, The hepatitis C virus non-structural protein 4B is an integral endoplasmic reticulum membrane protein, Virology 284 (2001) 70–81.
[11] M. Lundin, M. Monne, A. Widell, G. Von Heijne, M.A. Persson, Topology of the membrane-associated hepatitis C virus protein NS4B, J. Virol. 77 (2003) 5428–5438.
[12] D. Egger, B. Wolk, R. Gosert, L. Bianchi, H.E. Blum, D. Moradpour, K. Bienz, Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex, J. Virol. 76 (2002) 5974–5984.
[13] C. Lin, J.W. Wu, K. Hsiao, M.S. Su, The hepatitis C virus NS4A protein: interactions with the NS4B and NS5A proteins, J. Virol. 71 (1997) 6465–6471.
[14] L. Qu, L.K. McMullan, C.M. Rice, Isolation and characterization of non-cytopathic pestivirus mutants reveals a role for non-structural protein NS4B in viral cytopathogenicity, J. Virol. 75 (2001) 10651–10662.
[15] P. Neddermann, A. Clementi, R. De Francesco, Hyperphosphorylation of the hepatitis C virus NS5A protein requires an active NS3 protease, NS4A, NS4B, and NS5A encoded on the same polyprotein, J. Virol. 73 (1999) 9984–9991.
[16] N. Kato, H. Yoshida, S. Kioko Ono-Nita, J. Kato, T. Goto, M. Otsuka, K. Lan, K. Matsushima, Y. Shiratori, M. Omata, Activation of intracellular signaling by hepatitis B and C viruses: C-viral core is the most potent signal inducer, Hepatology 32 (2000) 405–412.
[17] H. Yoshida, N. Kato, Y. Shiratori, M. Otsuka, S. Maeda, J. Kato, M. Omata, Hepatitis C virus core protein activates nuclear factor kappa B-dependent signaling through tumor necrosis factor receptor-associated factor, J. Biol. Chem. 276 (2001) 16399–16405.
[18] S. Einav, M. Elazar, T. Danieli, J.S. Glenn, A nucleotide binding motif in hepatitis C virus (HCV) NS4B mediates HCV RNA replication, J. Virol. 78 (2004) 11288–11295.
[19] A. Bateman, L. Coin, R. Durbin, R.D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E.L. Sonnhammer, D.J. Studholme, C. Yeats, S.R. Eddy, The Pfam protein families database, Nucleic Acids Res. 32 (2004) D138–D141.
[20] D.L. Wheeler, T. Barrett, D.A. Benson, S.H. Bryant, K. Canese, D.M. Church, M. DiCuccio, R. Edgar, S. Federhen, W. Helmberg, D.L. Kenton, O. Khovayko, D.J. Lipman, T.L. Madden, D.R. Maglott, J. Ostell, J.U. Pontius, K.D. Pruitt, G.D. Schuler, L.M. Schriml, E. Sequeira, S.T. Sherry, K. Sirotkin, G. Starchenko, T.O. Suzek, R. Tatusov, T.A. Tatusova, L. Wagner, E. Yaschenko, Database resources of the National Center for Biotechnology Information, Nucleic Acids Res. 33 (2005) D39–D45.
[21] A. Bairoch, R. Apweiler, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M.J. Martin, D.A. Natale, C. O'Donovan, N. Redaschi, L.S. Yeh, The universal protein resource (UniProt), Nucleic Acids Res. 33 (2005) D154–D159.
[22] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[23] N. Kato, M. Hijikata, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi, T. Sugimura, K. Shimotohno, Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 9524–9528.
[24] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, J.D. Thompson, Multiple sequence alignment with the Clustal series of programs, Nucleic Acids Res. 31 (2003) 3497–3500.
[25] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32 (2004) 1792–1797.
[26] N. Galtier, M. Gouy, C. Gautier, SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny, Comput. Appl. Biosci. 12 (1996) 543–548.
[27] K.B. Nicholas, H.B.J. Nicholas, GeneDoc: a tool for editing and annotating multiple sequence alignments, Distributed by the author, 1997.
[28] D. Juan, O. Grana, F. Pazos, P. Fariselli, R. Casadio, A. Valencia, A neural network approach to evaluate fold recognition results, Proteins 50 (2003) 600–608.
[29] G. Pollastri, D. Przybylski, B. Rost, P. Baldi, Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles, Proteins 47 (2002) 228–235.
[30] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (2000) 404–405.
[31] M. Albrecht, S.C. Tosatto, T. Lengauer, G. Valle, Simple consensus procedures are effective and sufficient in secondary structure prediction, Protein Eng. 16 (2003) 459–462.
[32] E.L. Sonnhammer, G. von Heijne, A. Krogh, A hidden Markov model for predicting transmembrane helices in protein sequences, Proc. Int. Conf. Intell. Syst. Mol. Biol. 6 (1998) 175–182.
[33] H. Zhou, Y. Zhou, Predicting the topology of transmembrane helical proteins using mean burial propensity and a hidden-Markov-model-based method, Protein Sci. 12 (2003) 1547–1555.
[34] G.E. Tusnady, I. Simon, The HMMTOP transmembrane topology prediction server, Bioinformatics 17 (2001) 849–850.
[35] R. Linding, L.J. Jensen, F. Diella, P. Bork, T.J. Gibson, R.B. Russell, Protein disorder prediction: implications for structural proteomics, Structure 11 (2003) 1453–1459.
[36] V. Neduva, R. Linding, I. Su-Angrand, A. Stark, F. de Masi, T.J. Gibson, J. Lewis, L. Serrano, R.B. Russell, Systematic discovery of new recognition peptides mediating protein interaction networks, PLoS Biol. 3 (2005) e405.
[37] A. Lupas, M. Van Dyke, J. Stock, Predicting coiled coils from protein sequences, Science 252 (1991) 1162–1164.
[38] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190.
[39] F. Abascal, R. Zardoya, D. Posada, ProtTest: selection of best-fit models of protein evolution, Bioinformatics 21 (2005) 2104–2105.
[40] A. Drummond, K. Strimmer, PAL: an object-oriented programming library for molecular evolution and phylogenetics, Bioinformatics 17 (2001) 662–663.
[41] S. Guindon, O. Gascuel, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, Syst. Biol. 52 (2003) 696–704.
[42] M.W. Dimmic, J.S. Rest, D.P. Mindell, R.A. Goldstein, rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny, J. Mol. Evol. 55 (2002) 65–73.
[43] T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learning: data mining, inference, and prediction, ed., Springer-Verlag, 2001.
[44] W.Y. Tong, M. Nagano-Fujii, R. Hidajat, L. Deng, Y. Takigawa, H. Hotta, Physical interaction between hepatitis C virus NS4B protein and CREB-RP/ATF6beta, Biochem. Biophys. Res. Commun. 299 (2002) 366–372.
[45] P. Burkhard, J. Stetefeld, S.V. Strelkov, Coiled coils: a highly versatile protein folding motif, Trends Cell Biol. 11 (2001) 82–88.
[46] P. Burkhard, M. Meier, A. Lustig, Design of a minimal protein oligomerization domain by a structural approach, Protein Sci. 9 (2000) 2294–2301.
[47] D. Krylov, J. Barchi, C. Vinson, Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids, J. Mol. Biol. 279 (1998) 959–972.
[48] J.R. Moll, M. Olive, C. Vinson, Attractive interhelical electrostatic interactions in the proline- and acidic-rich region (PAR) leucine zipper subfamily preclude heterodimerization with other basic leucine zipper subfamilies, J. Biol. Chem. 275 (2000) 34826–34832.
[49] E.J. Spek, A.H. Bui, M. Lu, N.R. Kallenbach, Surface salt bridges stabilize the GCN4 leucine zipper, Protein Sci. 7 (1998) 2431–2437.
[50] T.H. Lin, D.P. Grandgenett, Retrovirus integrase: identification of a potential leucine zipper motif, Protein Eng. 4 (1991) 435–441.
[51] G. Li, X.C. Zhang, GTP hydrolysis mechanism of Ras-like GTPases, J. Mol. Biol. 340 (2004) 921–932.
[52] I.R. Vetter, A. Wittinghofer, The guanine nucleotide-binding switch in three dimensions, Science 294 (2001) 1299–1304.
[53] J.M. Caruthers, D.B. McKay, Helicase structure and mechanism, Curr. Opin. Struct. Biol. 12 (2002) 123–133.
[54] M. Albrecht, F.S. Domingues, S. Schreiber, T. Lengauer, Structural localization of disease-associated sequence variations in the NACHT

and LRR domains of PYPAF1 and NOD2, FEBS Lett. 554 (2003) 520–528.
[55] G.Y. Yu, K.J. Lee, L. Gao, M.M. Lai, Palmitoylation and polymerization of hepatitis C virus NS4B protein, J. Virol. 80 (2006) 6013–6023.
[56] M. Elazar, P. Liu, C.M. Rice, J.S. Glenn, An N-terminal amphipathic helix in hepatitis C virus (HCV) NS4B mediates membrane association, correct localization of replication complex proteins, and HCV RNA replication, J. Virol. 78 (2004) 11393–11400.
[57] J. Ye, R.B. Rawson, R. Komuro, X. Chen, U.P. Dave, R. Prywes, M.S. Brown, J.L. Goldstein, ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs, Mol. Cell 6 (2000) 1355–1364.
[58] A.R. Ciccaglione, A. Costantino, E. Tritarelli, C. Marcantonio, M. Equestre, N. Marziliano, M. Rapicetta, Activation of endoplasmic reticulum stress response by hepatitis C virus proteins, Arch. Virol. (2005).
[59] K.D. Tardif, K. Mori, R.J. Kaufman, A. Siddiqui, Hepatitis C virus suppresses the IRE1-XBP1 pathway of the unfolded protein response, J. Biol. Chem. 279 (2004) 17158–17164.
[60] K.D. Tardif, K. Mori, A. Siddiqui, Hepatitis C virus subgenomic replicons induce endoplasmic reticulum stress activating an intracellular signaling pathway, J. Virol. 76 (2002) 7453–7459.
[61] K.D. Tardif, G. Waris, A. Siddiqui, Hepatitis C virus, ER stress, and oxidative stress, Trends Microbiol. 13 (2005) 159–163.
[62] G. Waris, K.D. Tardif, A. Siddiqui, Endoplasmic reticulum (ER) stress: hepatitis C virus induces an ER-nucleus signal transduction pathway and activates NF-kappaB and STAT-3, Biochem. Pharmacol. 64 (2002) 1425–1430.
[63] K. Namba, K. Naka, H. Dansako, A. Nozaki, M. Ikeda, Y. Shiratori, K. Shimotohno, N. Kato, Establishment of hepatitis C virus replicon cell lines possessing interferon-resistant phenotype, Biochem. Biophys. Res. Commun. 323 (2004) 299–309.
[64] N.J. Adams, R.W. Chamberlain, L.A. Taylor, F. Davidson, C.K. Lin, R.M. Elliott, P. Simmonds, Complete coding sequence of hepatitis C virus genotype 6a, Biochem. Biophys. Res. Commun. 234 (1997) 393–396.
[65] G. Inchauspe, S. Zebedee, D.H. Lee, M. Sugitani, M. Nasoff, A.M. Prince, Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 10292–10296.
[66] A. Takamizawa, C. Mori, I. Fuke, S. Manabe, S. Murakami, J. Fujita, E. Onishi, T. Andoh, I. Yoshida, H. Okayama, Structure and organization of the hepatitis C virus genome isolated from human carriers, J. Virol. 65 (1991) 1105–1113.
[67] T. Tanaka, N. Kato, M. Nakagawa, Y. Ootsuyama, M.J. Cho, T. Nakazawa, M. Hijikata, Y. Ishimura, K. Shimotohno, Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals, Virus Res. 23 (1992) 39–53.
[68] P.J. Chen, M.H. Lin, K.F. Tai, P.C. Liu, C.J. Lin, D.S. Chen, The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5′ termini of viral genomic and antigenomic RNA, Virology 188 (1992) 102–113.
[69] H. Okamoto, S. Okada, Y. Sugiyama, K. Kurai, H. Iizuka, A. Machida, Y. Miyakawa, M. Mayumi, Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions, J. Gen. Virol. 72 (Part 11) (1991) 2697–2704.
[70] H. Okamoto, K. Kurai, S. Okada, K. Yamamoto, H. Lizuka, T. Tanaka, S. Fukuda, F. Tsuda, S. Mishiro, Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes, Virology 188 (1992) 331–341.
[71] T.P. Leary, S.M. Desai, J. Yamaguchi, M.L. Chalmers, G.G. Schlauder, G.J. Dawson, I.K. Mushahwar, Species-specific variants of GB virus A in captive monkeys, J. Virol. 70 (1996) 9028–9030.
[72] J.N. Simons, T.J. Pilot-Matias, T.P. Leary, G.J. Dawson, S.M. Desai, G.G. Schlauder, A.S. Muerhoff, J.C. Erker, S.L. Buijk, M.L. Chalmers, et al., Identification of two flavivirus-like genomes in the GB hepatitis agent, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 3401–3405.
[73] J. Bukh, C.L. Apgar, M. Yanagi, Toward a surrogate model for hepatitis C virus: an infectious molecular clone of the GB virus-B hepatitis agent, Virology 262 (1999) 470–478.
[74] M. Kempfer, M.C. Nuebling, J. Loewer, Nucleotide sequence, submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
[75] B. Zhao, E. Mackow, A. Buckler-White, L. Markoff, R.M. Chancock, C.-J. Lai, Y. Makino, Cloning full-length dengue type 4 viral DNA sequences: analysis of genes coding for structural proteins, Virology 155 (1986) 77–88.
[76] G. Coia, M.D. Parker, G. Speight, M.E. Byrne, E.G. Westaway, Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene order and characteristics of the virus-specified proteins, J. Gen. Virol. 69 (Part 1) (1988) 1–21.
[77] C.W. Mandl, C. Kunz, F.X. Heinz, Presence of poly(A) in a flavivirus: significant differences between the 3′ non-coding regions of the genomic RNAs of tick-borne encephalitis virus strains, J. Virol. 65 (1991) 4070–4077.
[78] M.R. Pisano, V. Mercier, V. Deubel, H. Tolou, Complete nucleotide sequence and phylogeny of an American strain of yellow fever virus, TRINID79A, Arch. Virol. 144 (1999) 1837–1843.
[79] G. Aviles, J. Rowe, J. Meissner, J.C. Manzur Caffarena, D. Enria, S. St Jeor, Phylogenetic relationships of dengue-1 viruses from Argentina and Paraguay, Arch. Virol. 147 (2002) 2075–2087.
[80] J. Yang, P.Y. Yang, Z.J. Hu, E.D. Qin, M. Yu, W. Ou, Y.S. Zhou, Nucleotide sequence, submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
[81] C.N. Wu, Nucleotide sequence, submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
[82] G. Meyers, N. Tautz, P. Becher, H.J. Thiel, B.M. Kummerer, Recovery of cytopathogenic and non-cytopathogenic bovine viral diarrhea viruses from cDNA constructs, J. Virol. 70 (1996) 8606–8613.
[83] G. Meyers, T. Rumenapf, H.J. Thiel, Molecular cloning and nucleotide sequence of the genome of hog cholera virus, Virology 171 (1989) 555–567.