UNIVERSITY OF THE
FREE STATE
UNIVERSITEIT VAN DIE
VRYSTAAT
YUNIVESITHI YA
FREISTATA

PREPARATION OF RECOMBINANT ANTIGENS FOR DEMONSTRATING ANTIBODY RESPONSES IN PATIENTS WITH CRIMEAN-CONGO HAEMORRHAGIC FEVER VIRUS INFECTIONS

Rudo Ruth Samudzi

2011

University of the Free State

Bloemfontein Campus

UNIVERSITY OF THE
FREE STATE
UNIVERSITEIT VAN DIE
VRYSTAAT
YUNIVESITHI YA
FREISTATA

UFS
UV

PREPARATION OF RECOMBINANT ANTIGENS FOR DEMONSTRATING ANTIBODY RESPONSES IN PATIENTS WITH CRIMEAN-CONGO HAEMORRHAGIC FEVER VIRUS INFECTIONS

Rudo Ruth Samudzi

B.Med.Sc

Dissertation submitted in fulfilment of the requirements for the degree Master of Medical Science at the University of the Free State

Promoter: **Prof F.J. Burt**

Department of Medical Mircobiology and Virology,
Faculty of Health Sciences, University of the Free State

University of the Free State, Bloemfontein Campus

June 2011

TABLE OF CONTENTS

DECLARATION

I declare that the dissertation hereby submitted by me for the M.Med.Sc (Virology) degree at the University of the Free State, Bloemfontein, is my own independent work and has not been previously submitted by me at another institution/faculty. I further more cede copyright in favour of the University of the Free State.

_____

Rudo Ruth Samudzi

08/06/2011

ACKNOWLEDGEMENTS

I would like to thank the following persons and institutions:

- Prof Felicity. J. Burt for her wonderful work in supervising this project. Thank you for your willingness to help me whenever I encountered problems. Thank you for your encouragement and scientific knowledge that added quality to the project. I have really learned a lot and am looking forward to learning more.

- The Department of Medical Microbiology and Virology, Faculty of Health Sciences for providing the facilities that enabled me to complete this project.

- The Polio Research Foundation for financial assistance throughout the period of my study.

- My colleagues and friends for support not only academically but also on a personal level. Thank you to Lehlohonolo Mathengtheng, Kulsum Kondiah, Shannon Smouse, Mitta Mamabolo, Manie Hanekom, special thanks to Azeeza Rangunwala for lending me your laptop to type up, to Carina Combrinck and Arina Jansen for assistance with editing, my childhood friends Beatrice Kyambadde, Eriva Kyambadde, Grace Mukasa, Rita Buanyomi for their support. To all those I have not mentioned, you're not forgotten. Thank you.

- My parents and siblings for their prayers and faith in me. I will always appreciate you.

- My boyfriend, Vernon Mwazi for always being there for me even at the most stressful moments. Your motivation, love and faith in me has made this possible.

- God, my Saviour and Lord who has carried me through this challenging time and blessed me in all aspects of my life.

ABSTRACT

Crimean-Congo haemorrhagic fever (CCHF) is a tick-borne viral zoonosis widely distributed in Africa, Asia, Russia and the Balkans. The causative agent, CCHF virus (CCHFV) has the propensity to cause nosocomial infections with a high fatality rate. Cases of CCHF are diagnosed annually in southern Africa. Increasing numbers of cases are seen in regions of Asia and in the past ten years CCHFV has emerged in several countries in the Balkans and re-emergence in south-western regions of the Russian Federation. Diagnosis of CCHFV infections during the acute phase is based on isolation of the virus or amplification of viral RNA. Patients that survive the infection have a demonstrable IgG and IgM antibody response, usually from day 5 to 7 after onset of illness. Current serological diagnostic assays based on ELISA or IF use inactivated virus which requires biosafety level 4 facilities for culturing the virus and therefore limits the number of laboratories that can prepare suitable reagents. Preparation of recombinant antigens would enable laboratories to perform serological diagnosis of CCHFV infections and surveillance studies. The purpose of this study was to prepare a recombinant CCHFV nucleoprotein using a bacterial expression system, to determine if the protein was immunogenic and to determine if the protein was able to detect IgG antibodies in survivors of CCHFV infection.

The complete open reading frame of the gene encoding the NP of CCHFV was amplified by RT-PCR using primers specifically designed with restriction sites engineered to the primers to facilitate cloning. The amplicon was cloned into pGEM® T Easy vector using T/A cloning and the gene sequenced to confirm that the correct gene had been amplified and cloned into the vector for downstream cloning and expression applications. Initially we aimed to express the native gene using a bacterial expression system and the NP gene was rescued from the recombinant plasmid and cloned into pQE-80L vector using the BamH1 and Pst1 restriction sites present in the multiple cloning site on the vector. Various attempts were made to express the CCHFV NP protein however no protein was detectable using SDS PAGE methods or Western blot.

The nucleotide sequence that we had determined for the open reading frame of our gene encoding the NP was analysed using the Rare Codon Analysis Tool software and we elected to codon optimize the gene for expression in *E. coli*. The optimized gene was synthesized by GenScript and supplied cloned in the multiple cloning site of pUC57. The optimized gene was excised from pUC57 and cloned into pColdTF bacterial expression vector. A 106 kDa protein was expressed from the construct likely representing the HIS tagged TF chaperone protein fused to the CCHFV NP protein and confirmed by Western blot analysis. A higher yield of the protein was present in the insoluble phase and as optimization of the growth and induction conditions did not significantly alter the insoluble to soluble ratio of the expressed protein, the protein was harvested from the insoluble phase by denaturing, purification and refolding of the protein. The biological activity of the recombinant protein was confirmed using immunoassays and by immunizing mice to determine if the antibodies induced by the recombinant protein could be detected using an antigen prepared from the whole virus. Four of five mice immunized with the recombinant NP had a detectable antibody response using an immunofluorescent assay. Serum samples from acute and convalescent patients collected at varying stages after onset of illness were reacted in a Western blot with the recombinant CCHFV NP protein. The recombinant antigen was able to detect IgG antibody in all the convalescent patient sera except two sera collected on days 14 and 15 during the acute phase. In contrast all the samples were detected using the recombinant antigen in an ELISA. Due to the potential biohazardous nature of samples only samples collected two weeks after onset of illness were tested. The results showed 100% concordance with the results obtained in an ELISA using mouse brain derived antigen. The assay was shown to be reproducible and stability studies showed that four months after preparation the protein was still active. A full validation of the protein using a large panel of serum samples from confirmed CCHF patients is now required.

The results suggest that bacterially expressed proteins lacking post translational modifications and folding that occur with mammalian and baculovirus expression can be used in ELISA to detect IgG antibody against CCHFV in human sera which finds application in diagnostics, epidemiologic and surveillance studies.

CHAPTER 1

LITERATURE REVIEW

1.1 Introduction and history

Crimean-Congo haemorrhagic fever virus (CCHFV) is a tick-borne arbovirus that occurs in Africa, Asia, Russia and the Balkans (Hoogstraal, 1979; Maltezou et al., 2010). Humans acquire CCHFV infection after being bitten by an infected tick or through direct contact with the blood and tissues of infected livestock (Swanepoel et al., 1989).

CCHFV was first described in people bitten by ticks in 1944 during a large outbreak of haemorrhagic fever that occurred among military workers in the Crimean Peninsula, Russia (Hoogstraal, 1979). The disease was then known as Crimean haemorrhagic fever (CHF). The following year human subjects were inoculated with the virus and it was shown that the disease was caused by a filterable pathogen present in the blood of patients during the acute phase of illness. It was also shown that the pathogen was present in suspensions prepared from ticks, which were suspected to be vectors of the pathogen. In 1967, the virus was first propagated in new born white mice and the antigenic, physiologic and morphologic characteristics of the virus were established (Casals, 1969). In 1969, Crimean hemorrhagic fever virus was found to be antigenically similar to Congo virus isolated in 1956 from a febrile patient in the Democratic Republic of Congo, and thereafter named Crimean-Congo haemorrhagic fever virus (Casals, 1969; Chumakov et al., 1970)

1.2 Virus classification, characteristics and biology

CCHFV is a member of the *Nairovirus* genus of the *Bunyaviridae* family (Casals and Tignor, 1980; Calisher and Karabatsos, 1989). The genus *Nairovirus* consists of 34 viruses that are divided into seven different serogroups based on antigenic relationships. CCHFV, Hazara virus from Pakistan and Khasan virus from the former Union of Soviet Socialist Republics (USSR) all form part of the CCHFV serogroup. Apart from CCHFV, the only other members of the genus that are the

cause of significant human disease are Nairobi sheep disease and Dugbe viruses. Nairobi sheep disease virus which comes from East Africa is a tick-borne pathogen of sheep and goats and sporadically causes mild disease in humans (Davies et al., 1978). Dugbe virus is a tick-borne virus from West Africa known to cause mild infection in cattle and sheep and less frequently causes mild illness in humans (Burt et al., 1996). Nairoviruses were classified initially based on antigenic relationships, however the groupings have subsequently been verified through demonstration of morphological and molecular similarities among the viruses (Calisher & Karabatsos, 1989).

The virions of members of the *Bunyaviridae* family are spherical in shape with diameters of between 90-100nm (Clerx et al., 1981). As shown in the schematic diagram in Figure 1, the single-stranded, negative-sense RNA genome consists of an S segment that encodes the nucleocapsid, an M segment that encodes two envelope glycoproteins G_n and G_c and an L segment that encodes the RNA dependent RNA polymerase (Clerx et al., 1981; Elliott, 1990).

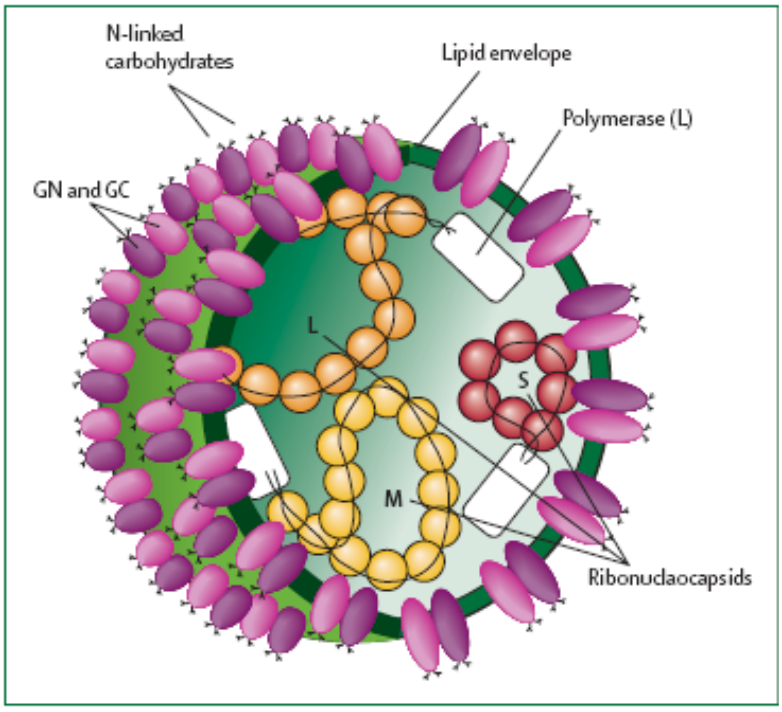


Figure 1. Schematic presentation of the virus structure (Ergonul, 2006).

Few studies have been done on the replication strategies of CCHFV. The S segment of Dugbe virus, the most broadly studied member of this genus, has one

open reading frame encoding the nucleoprotein (NP) of the virus (Ward et al., 1990; Sanchez et al., 2002). The predicted amino acid sequences from the S segment of CCHFV and Hazara viruses were aligned with the predicted amino acid sequence of Dugbe virus and significant sequence homology was shown among the three viruses confirming that the S segment of CCHFV and Hazara encode the viral NP. CCHFV S RNA comprises approximately 1672 nucleotides and has a single open reading frame which encodes the nucleoprotein (54 x 10^3Da), the major structural protein of the virus (Marriott and Nuttall, 1992). Within the *Bunyaviridae* family the NP has been shown to be the most abundant and immunodominant viral protein and therefore induces high levels of specific humoral antibodies (Magurano and Nicoletti, 1999). Laboratory diagnostic reagents have been prepared from recombinant NP (Marriott et al., 1994; Saijo et al., 2002a, 2002b; Saijo et al., 2005, Garcia et al., 2006).

The M segment, which is approximately 5367 nucleotides in length, has one open reading frame which encodes for a precursor polypeptide that is post-translationally cleaved into mature G_n and G_c glycoproteins (Sanchez et al., 2002). Analyses performed on the polyprotein sequence of a CCHFV strain indicate that the tetrapeptide RRLL is the 5' cleavage site for mature G_n and tetrapeptide RKPL is the 5' cleavage site for mature G_c site as shown in Figure 2. Similarly, tetrapeptide RKLL is assumed to be the site of cleavage for the 3' end of G_n. These tetrapeptides are conserved in the M segment sequences of all the CCHFV strains that have been analysed. The predicted sizes of the G_n and G_c are 37kDa and 75kDa respectively (Sanchez et al., 2002; Papa et al., 2002).

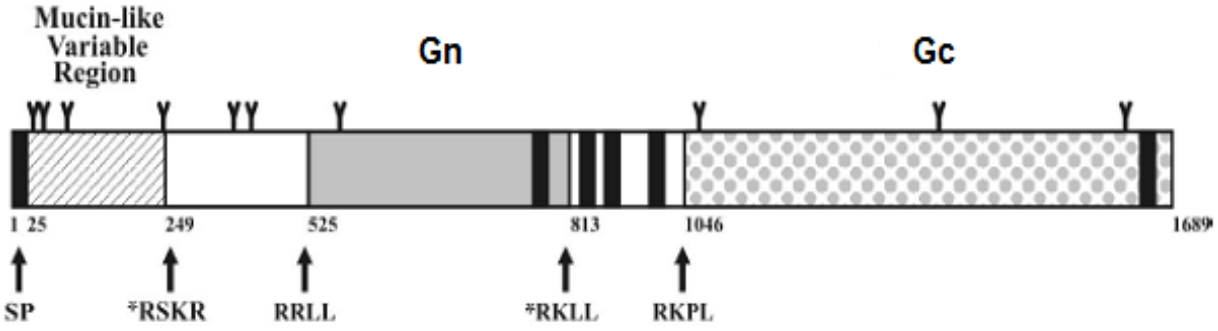


Figure 2. Schematic representation of the glycoprotein open reading frame of CCHFV Matin strain (Sanchez et al., 2002)

By analogy of CCHFV with other bunyaviruses the L segment encodes a large protein namely the RNA-dependent RNA polymerase estimated to be between 12000-14000 nucleotides (Marriott & Nuttall, 1996; Deyde et al., 2006).

Although the complete pathogenesis of CCHFV is not fully understood, the glycoproteins are thought to influence the vertebrate host range and cell tropism of the virus and are the sites where neutralizing antibodies bind (Ahmed et al., 2005). By analogy with other members of the *Bunyaviridae* family, CCHFV attaches to host cell receptors via the glycoproteins. Although the cellular receptor for CCHFV has not yet been defined, the virus has been suggested to enter the host by receptor-mediated endocytosis and replicates in the cytoplasm (Schmaljohn and Patterson 1990; Whitehouse, 2004). Genetic studies have demonstrated high levels of diversity among the CCHFV isolates from geographically distinct areas (Casals, 1969; Marriott & Nuttall, 1996; Rodriguez et al., 1997; Papa et al., 2002; Hewson et al., 2004a; Burt and Swanepoel, 2005; Deyde et al., 2006). The viral genome is significantly variable with 20%, 31% and 22% nucleotide differences reported for the S, M and L genes respectively (Deyde et al., 2006, Burt and Swanepoel, 2005; Hewson et al., 2004a). However despite these high levels of heterogeneity, the isolates seem to comprise a single virus species (Casals, et al., 1969).

Many of these nucleotide differences are synonymous resulting in only 8% diversity within the deduced amino acid sequence for the NP. This may explain the antigenic similarity between isolates. Phylogenetic analyses of partial and complete nucleotide sequence data for the S segment have identified distinct S segment genotypes related to geographical distribution and subsequently designated Asia 1 and 2, Africa 1, 2 and 3 and Europe 1 and 2 (Hewson et al., 2004a). Genetic studies have also shown that similar genotypes can be geographically distinct. The mechanism for the dispersal of ticks and subsequently virus are likely to include the movement of migrating birds annually from Europe and Asia to Africa (Hoogstraal et al., 1961). Ticks can also be disseminated between continents by movement of livestock during trade. However, it also appears from phylogenetic analyses of CCHFV isolates that virus circulation is grouped mainly within the two land masses of Africa and Eurasia where the distribution of the strains of the virus within the

continents is most likely proportional to the distribution and spread of the vectors of the virus (Marriot & Nuttall, 1996; Burt et al., 1996; Papa et al., 2002).

There is evidence for the occurrence of reassortment in nature and recombination events, although probably less significant for CCHFV, have been shown to occur (Deyde et al., 2006, Hewson et al., 2004b, Kondiah et al., 2010). Phylogeny based on S and L sequence data show greater similarity compared to phylogeny based on M segment data. Incongruencies in groupings of isolates provide evidence for reassortment of genes (Burt et al., 2009; Hewson et al., 2004b). There appears to be a higher frequency of reassortment associated with the M segment however this could also be a reflection of the fitness of viable reassorted virus. Nonetheless the genetic diversity and occurrence of reassortment need to be considered when designing molecular or recombinant diagnostic tools.

Little is known about the stability of CCHFV, but as it is an enveloped virus, it is sensitive to lipid solvents (Karabatsos, 1985) and is inactivated by low concentrations of formalin and beta-propriolactone. The virus is unstable in infected human tissues after death (Hoogstraal, 1979 ; Butenko and Chumakov, 1990), but the analysis of specimens from human patients seem to show that the virus is preserved for at least a few days at ambient temperature in separated serum. Infectivity is destroyed by boiling or autoclaving, but the virus remains stable at temperatures -60ºC (Hoogstraal, 1979; Clerx et al., 1981; Watts et al., 1989b). CCHFV replicates to low titers in a number of cell lines such as Vero cells, chicken embryo related (CER) cells and baby hamster kidney 21 cells (BHK-21). Strains of CCHFV differ in their ability to replicate and produce plaques in different cell lines. The Human Adrenal Gland Adenocarcinoma (SW-13) and CER cells have been used for plaque assays, (Watts et al., 1989b) with plaque formation visualized using an appropriate dye such as neutral red (Shepherd et al., 1986). Virus isolation and titration is mainly done by intracerebral inoculation of day-old mice (Hoogstraal, 1979).

1.3 CCHFV reservoirs and vectors

CCHFV has been isolated from a minimum of 31 species of ticks of 7 genera,

including 29 ixodids (hard-bodied) and 2 argasids (soft-bodied). For the majority of the species there is no evidence that they are vectors of the disease, and in some cases the virus recovered from engorged ticks may have been present in the blood meal obtained from a viremic host (Hoogstraal, 1979; Camicas et al., 1991). Argasid ticks are not likely to transmit infection since CCHFV failed to replicate in four species namely *Argas walkerae, Ornithodorus porcinus, O. savigny* and *O. sonrai*, following attempts to infect them experimentally. (Shepherd et al., 1989a; Durden et al., 1993). However, it has been shown that the ixodid ticks of several genera can serve as vectors for CCHFV infection. Transstadial and transovarial transmission of CCHFV from adult females to larvae has also been demonstrated in a few members of the *Hyalomma, Dermacentor, and Rhipicephalus* genera (Kondratenko, 1976). However, the correlation in the distribution of CCHFV and the distribution of the *Hyalomma* ticks strongly indicates that members of this genus are the principle vectors of the virus (Hoogstraal, 1979; Shepherd et al., 1991).

Ixodid ticks have three development stages in their life cycle, larvae, nymph and adults, each of which attaches to vertebrate hosts to feed before molting to the next stage and feeding again (Hoogstraal, 1979). Many *Hyalommas* are two-host ticks and the transmission cycle is illustrated in Figure 3. The adult females drop off the host to lay eggs. The eggs then hatch into six-legged larvae which become nymphs and feed on the small vertebrates such as rodent, hedgehogs and hares. The engorged nymphs then drop off these hosts and molt into adults which feed on large mammals such as sheep, cattle and goats. Transmission to humans most frequently occurs when humans handle the infected tissues of these large mammals or from the bite of an infected tick (Kondratenko, 1976). Infected humans are then hospitalized and the possibility of human to human transmission and nosocomial outbreaks can occur mainly among health care workers (Ergonul, 2006).

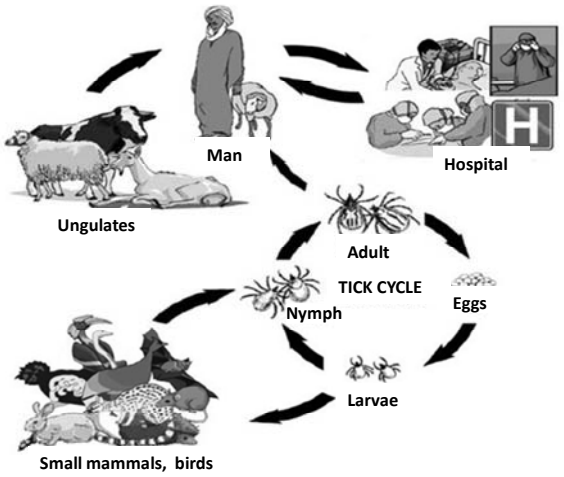


Figure 3. Schematic representation of two-host tick cycle of *Hyalomma* species and the natural cycle of CCHFV (http://www.dpd.cdc.gov/dpdx).

The virus causes a transient viremia which has been demonstrated in most small mammals of Europe, Asia and Africa, and in some cases it has been shown that these hosts can transmit the virus to ticks. Large mammals also develop detectable levels of viremia and can infect ticks (Shepherd et al. 1989a; Shepherd et al. 1991). The sera of wild vertebrates in South Africa and Zimbabwe have a low antibody prevalence to CCHFV, but the prevalence is highest in large herbivores such as the zebra, buffalo, rhinoceros and giraffe, which are the preferred hosts of the adult *Hyalomma* ticks (Shepherd et al., 1987; Burt et al., 1993). Adult *Hyalomma* ticks can feed on ostriches and antibody has also been found in their sera, but not in the sera of wild passerines or water birds. (Shepherd et al., 1987). Immature *Hyalomma* ticks feed on small mammals and ground- frequenting birds and among these, antibody prevalence is highest in hares and also in a small proportion of rodents and guinea fowl (Shepherd et al., 1987). Despite the fact that antibody surveys have demonstrated that high rates of infection occur in livestock, it cannot be concluded that large vertebrates maintain CCHFV because they are hosts to adult *Hyalommas* in which transovarial transmission occurs minimally. It was therefore hypothesized that the cycle of infection between immature ticks and small vertebrates supposedly constitutes the most significant amplifying mechanism which ensures maintenance of the virus and promotes transstadial transmission of infection by adult ticks to large mammals (Watts et al., 1989a).

Humans acquire CCHFV infection commonly from the bite of an infected tick, or

from direct contact with the blood and tissues of infected livestock. Most susceptible patients are adult males involved in the livestock industry, such as farmers, labourers, slaughtermen and veterinarians (Swanepoel et al., 1987; Watts et al., 1989a). Other patients who live in urban areas become susceptible to infection after contact with animal tissues or after being bitten by ticks during hunting or hiking trips. It is possible for humans to acquire CCHFV infection by simply squashing ticks between their fingers (Hoogstraal, 1979; Swanepoel et al., 1987). The occurrence of disease in slaughtermen suggests that sometimes viremic animals arrive at abattoirs. Infection within abattoirs occurs mainly in people involved in bleeding the animals at the initial stages of the slaughtering process as well as those who handle animal skins where semi-engorged ticks that detach from slaughtered animals attach to the closest available host (Swanepoel et al., 1998). Therefore infection seems to be limited to those who have come into contact with fresh blood or other tissues as infectivity may be destroyed by the drop in pH which occurs in tissues after death (Swanepoel et al., 1998).

The low incidence of infection occurring in humans can be attributed to the fact that viremia in livestock is of short duration with low levels of infectivity as compared to that in other zoonotic diseases such as Rift Valley fever, which is more readily acquired from contact with infected tissues. Even though a high proportion of patients become infected from tick bites, humans are not the preferred hosts of *Hyalomma* ticks and are not bitten as often as livestock.

1.4 Epidemiology

The distribution of CCHFV correlates directly with the distribution of the principle vectors of the virus, ticks of the *Hyalomma* genus. Incidences of naturally acquired human infection have been recorded in many different countries in Africa, Asia, Russia and the Balkans. These include China, Yugoslavia, Bulgaria, Albania, Kosovo, Turkey, Iraq, Iran, Pakistan, United Arab Emirates, Saudi Arabia, Oman, Tanzania, Central African Republic, Democratic Republic of Congo (formerly Zaire), Uganda, Kenya, Mauritania, Burkina Faso, South Africa, Namibia and most recently the virus has been identified as the aetiologic agent of disease in Greece (Hoogstraal, 1979; Burney et al., 1980; Suleiman et al., 1980; Al Tikriti et al., 1981;

Gear et al., 1982; Saluzzo et al., 1985; Schwarz et al., 1995; El Azazy & Scrimgeour, 1997; Papa et al., 2002; Dunster et al., 2002; Karti et al., 2004; Morikawa et al., 2007; Zavitsanou et al., 2009; Maltezou et al., 2010; Papa et al., 2010). The worldwide geographic distribution of CCHF viral isolates, human disease and distribution of the *Hylamma* species of ticks is illustrated in Figure 4.

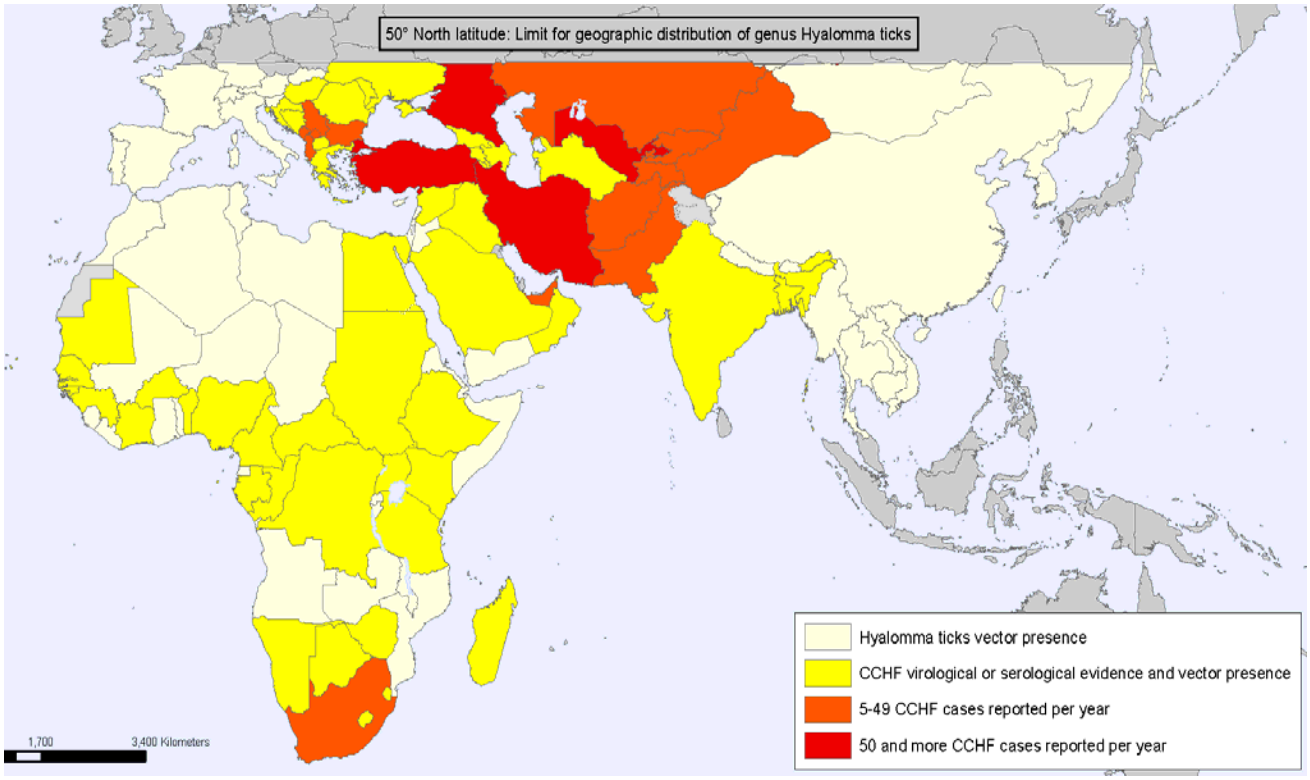


Figure 4. The geographic distribution of CCHF viral isolates and human disease (http://www.glews.net/images/Photos/CCHF_Risk_WHO.png).

The first outbreaks of CCHF described in the Crimean Peninsula, Russia in 1944 and 1945 happened during World War II where large numbers of soldiers and peasant farmers were bitten by ticks while harvesting crops and sleeping outdoors (Hoogstraal, 1979). Disease endemicity was recognized in many countries in eastern Europe and Asia through the occurrence of highly visible epidemics or nosocomial outbreaks creating opportunities for human intervention and resulting in people constantly being exposed to infection. These include the establishment of major land repossession plans or sudden changes in animal husbandry practices in the former Soviet Union and Bulgaria in the 1950s and 1960s, and in Rostov Province, Russia in 1999 (Hoogstraal, 1979).

Nosocomial outbreaks of infection also occurred in Pakistan in 1976 and in Dubai and Iraq in 1979. People were also exposed on many occasions to blood and ticks from handling and slaughtering livestock imported from Africa and Asia to Saudi Arabia in 1990, the United Arab Emirates in 1994-95 and Oman in 1995 (Al Tikriti et al., 1981; El Azazy & Scrimgeour, 1997; Hoogstraal, 1979). The perception that CCHF is an emerging disease arose as a result of the occurrence of these epidemics.

In many other countries in Africa, Asia, Russia and the Balkans (Morikawa et al., 2007; Zavitsanou et al., 2009), however, the presence of the virus was determined due to prospective laboratory examinations that were performed and not because a specific clinical agent had been identified. Serological surveys demonstrate that there is evidence of widespread circulation of CCHFV in nature in many different countries that still have to detect the occurrence of human disease (Hoogstraal, 1979).

In Africa, for example, there were only 15 human cases recognized up to 1979, 8 of which occurred among laboratory personnel (Swanepoel et al. 1987). Circumstances changed after identification of the first case of disease in South Africa in February 1981 whereby a child was bitten by a *Hyalomma* tick. More than 180 laboratory confirmed cases of CCHFV infection have been identified in southern Africa from 1981 - 2010 (NICD surveillance bulletins).

It was found that antibody to CCHFV was diversely distributed in the sera of livestock and wild vertebrates in South Africa, Zimbabwe and Namibia, together with sera that had been in frozen storage since 1964 (Shepherd et al., 1987; Burt et al., 1993). This suggests that the virus must have been circulating in southern Africa long before its presence was identified. It is therefore considered that regular diagnosis of CCHF infections in the subcontinent in recent years can be attributed to the increased awareness among medical clinicians, a consequence of the wide publicity given to the disease and due to the availability of a meticulous diagnostic service. Human disease has also been seen in Mauritania, Burkina Faso and Kenya implicating the presence of the virus throughout Africa and that disease severity in Africa is similar to that which occurs in eastern Europe and Asia

(Saluzzo et al., 1985; Dunster et al., 2002).

In recent years the virus has emerged in various Balkan countries and re-emerged in south western regions of the Russian Federation after a 27 year absence (Kuhn et al., 2004; Vorou et al., 2007; Maltezou and Papa, 2010). The reemergence is likely the result of changes in agricultural practice and land use as seen in 1944 in the Crimean peninsula. However in addition it is likely that changes in weather patterns that influence tick activity and populations now also play a role in the emergence or re-emergence of this virus (Randolph and Rigers, 2007). There is a concern that this virus could expand its current geographic distribution and become endemic in new regions.

1.5 Clinical manifestations

CCHFV infection in humans usually results in severe hemorrhagic disease. It has been observed that the course of CCHFV infection occurs through four distinct phases, namely incubation, pre-hemorrhagic, hemorrhagic and convalescent (Ergonul, 2006). The incubation period of CCHFV after a tick bite is between 1-3 days, however this can extend to 7 days, and is usually 5-6 days in people exposed to blood and tissues of infected livestock or of human patient (Swanepoel et al., 1987; Swanepoel et al., 1989). The pre-hemorrhagic phase is characterized by a sudden onset of headache, fever, chills, dizziness and photophobia. Nausea, vomiting and sore throat commonly occur and patients may experience non-localized abdominal pain as well as diarrhea. Fever often fluctuates and patients undergo mood changes over the next 2 days with feelings of confusion and aggression. By day 2–4 of illness patients feel loss of energy, depression, drowsiness and may have a flushed appearance with swelling of the conjunctivae. At this time, tenderness may be localized in the right upper quadrant of the abdomen and enlargement of the liver may be seen. There may be tachycardia with mild hypotension and lymphadenopathy, exanthema and petechiae of the throat, tonsils and buccal mucosa (Hoogstraal, 1979; Whitehouse, 2004; Ergonul, 2006). Haemorrhagic manifestations occur in severe cases 3-6 days after infection and range from petechial rash on the trunk and limbs and ecchymosis on the skin and mucous membranes as shown in Figure 5, to intestinal hemorrhage and bleeding

from other openings such as the mouth.

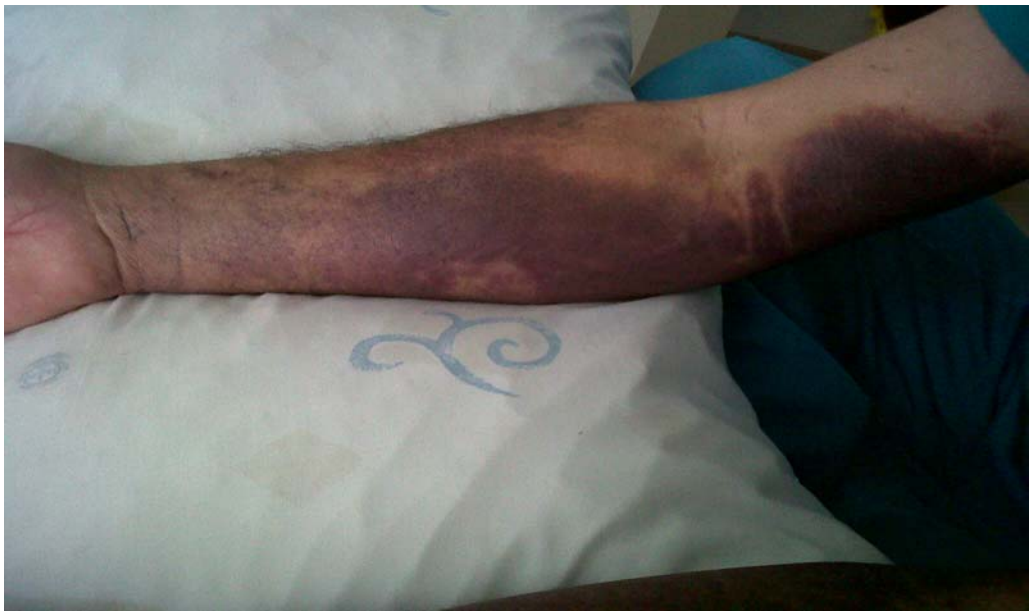


Figure 5. Massive cutaneous ecchymosis on the arm of a CCHF patient (Source Burt FJ, permission granted to use for academic purposes).

The mortality rate is approximately 30% with deaths occurring 5–14 days after onset of illness. In some patients progression to haemorrhagic disease does not occur or is less severe and thus convalescence takes place 15-20 days post infection. During this period patients generally feel weak, dizzy, nauseous and may experience loss of hearing or loss of memory and although these problems may persist for sometime, they are seldom permanent (Swanepoel et al.,1987; Swanepoel et al., 1989; Ergonul, 2006).

1.6 Clinical pathology and pathogenesis

During the first few days of illness there may be changes in the cellular and chemical composition of the human patients blood including leukocytosis or leucopenia and elevated levels of aspartate transaminase (AST), alanine transaminase (ALT), gamma-glutamyltransferase, lactic dehydrogenase, alkaline phosphatase and creatine kinase. Levels of bilirubin, creatinine and urea increase while serum protein levels decrease during the second week of illness (Joubert et al., 1985; Swanepoel et al., 1987; Swanepoel et al., 1989). There is also evidence of thrombocytopenia, fibrin degradation products as well as depression of fibrinogen and hemoglobin values during the first few days of illness.

Complete autopsies are rarely performed on patients that die of CCHF. Liver samples are usually taken with biopsy needles in order to examine tissues. Lesions found in the liver vary from disseminated areas of necrosis to very high levels of necrosis, involving 75% of hepatocytes, and haemorrhage (Baskerville et al., 1981; Burt et al., 1997; Joubert et al., 1985). Areas where necrosis occurs are normally marked by haemorrhage and cell loss linked to eosinophilic changes in hepatocytes (changes in hepatocytes observed by changes to cell staining).

Lesions present in the other organs include congestion, haemorrhage and cell death in the central nervous system, kidneys and adrenals and general loss of lymphoid tissues. None of the histopathologic features are diagnostic of CCHF as similar features can be observed in other viral, rickettsial and bacterial infections as well as toxic exposures. Therefore, a specific diagnosis can only be confirmed by immunohistochemical or virological tests (Burt et al., 1997).

The pathogenesis of CCHFV is not completely understood (Burt et al., 1997), but by analogy with other arthropod-borne virus infections it can be deduced that CCHFV may undergo some replication at the site of inoculation. There is also speculation that haematogeneous and lymph-borne spread of infection to organs such as the liver, the main sites of replication may occur. Using immunohistochemistry to localize CCHFV in tissues it has been revealed that mononuclear phagocytes and endothelial cells are also major targets for of virus infection (Burt et al., 1997). A large number of fatal haemorrhagic fever viruses also display similar tropism. The mononuclear phagocyte system may comprise a mechanism that neutralizes the virus in some patients, however in other patients the virus may multiply in these cells thus increasing levels of viremia. When mononuclear phagocytes are infected and lymphoid cells destroyed, the virus may be protected from phagocytosis allowing it to spread further. Also infection of mononuclear phagocytes and endothelial cells may contribute to the pathogenesis of CCHF via release of physiologically active substances, such as cytokines, tumour necrosis factor (TNF) and other inflammatory mediators and procoagulants (Swanepoel et al., 1989). Disseminated intravascular coagulopathy (DIC) seems to occur early and is vital in the pathogenesis of the disease. Hepatocytes are a major target of the virus and it

has been suggested that hepatocellur necrosis may be intervened by a direct viral cytopathic effect due to the occurrence of minor inflammatory infiltration. Hepatocellular necrosis causes further release of TNF and other procoagulants into the bloodstream, and eventually inhibits the production of coagulation factors to replace those utilized in DIC (Burt et al.,1997).

1.7 Laboratory diagnosis of CCHF

It is very important to make an early diagnosis of CCHFV with regards to providing supportive therapy, treating infected patients and preventing hospital acquired (nosocomial) infections (Donets et al., 1982, Shepherd et al., 1986, Shepherd et al., 1988). History of tick bites or exposure of patients to the blood and tissues of infected livestock are indicators of CCHFV infection (Hoogstraal, 1979). CCHFV is generally diagnosed either by virus isolation in tissue culture or suckling mice, by utilizing immunological assays such as ELISA for detection of IgG and IgM antibody response or by molecular methods such as reverse transcriptase polymerase chain reaction (RT-PCR) for detection of viral nucleic acid (Shepherd et al., 1986).

Diagnosis of CCHF by virus isolation and culturing should be performed in biosafety level (BSL) 4 containment facilities because CCHFV is highly pathogenic to humans. The classical method for isolation of CCHFV is by inoculating suckling mice intracranially with blood from an acute phase patient (Shepherd et al., 1986). Virus isolation in cell culture is easier and more rapid, but the sensitivity is not good and virus can only be detected when present in high concentrations (Shepherd et al 1986) However, the virus can be isolated from blood and organ suspensions in various susceptible cell lines including Rhesus Monkey Kidney Epithelial cells (LLC-MK2), Vero (monkey kidney cells), BHK-21, and SW-13 cells with maximal virus yields (10^7–10^8 plaque-forming units/ml) 4–7 days post incubation. The virus may produce little or no cytopathic effect (CPE) depending on the cell line and strain, and develop into a noncytopathic persistent infection of the cells. In this case the virus can be identified by performing an IFA with specific monoclonal antibodies. In addition, CPE and plaque formation may be observed only after serial passage of virus (Shepherd et al., 1986).

A number of serological assays have been used to diagnose CCHF infection such as complement fixation (CF), hemagglutination inhibition (HI), immunodiffussion, immunofluorescence assays (IFA) and enzyme linked immunosorbent assay (ELISA) (Casals and Tignor, 1980; Donets et al., 1982, Swanepoel et al., 1983; Shepherd et al., 1986, Shepherd et al., 1988, Shepherd et al., 1989b; Swanepoel et al., 1989; Burt et al., 1994). Previous studies however have demonstrated the lack of sensitivity and reproducibility of CF, HI and immunodiffusion assays. The use of IFA and ELISA have proven to be more reliable and useful in serological diagnosis as they are able to detect and differentiate IgM and IgG antibodies making recent and past infections easily distinguishable from one another (Shepherd et al., 1989b). The presene of IgM antibodies indicates a recent infection and IgG a past infection. IgM and IgG antibodies can both be detected by IFA approximately 7 days after onset of illness and are present in all survivors by the 9th day after onset of illness. The amount of IgM antibody decreases and is no longer detectable after 4 months, whereas even though IgG titers decrease gradually, they are still detectable for at least 5 years (Shepherd et al., 1989b, Burt et al., 1994). In fatal cases antibody is frequently not detectable. Recent or current infections can also be confirmed by demonstrating seroconversion, or a fourfold or greater rise in antibody titer in paired serum samples, or IgM antibody in a single sample. In fatal cases antibodies are very seldom detectable and diagnosis of CCHF infection is usually confirmed by virus isolation from the serum or liver biopsy specimens. There have been a limited number of studies in which recombinant CCHFV NPs have been developed and used in an ELISA (Saijo et al., 2002a; Tang et al., 2003) or in a IFA (Saijo et al., 2002b) to detect serum antibodies from infected patients. However, these assays have not been fully validated for use as diagnostic tools.

The use of molecular-based diagnostic assays, such as the reverse transcription-polymerase chain reaction (RT-PCR) provides many advantages and now often serves as the most utilized tool in the diagnosis of CCHF, and other viral haemorrhagic fevers (Drosten et al., 2003). RT-PCR can detect the nucleic acid of the virus and can be designed to be highly specific, therefore a final diagnosis of CCHF can be made without the need to culture the virus, which would require the use of specialized BSL 4 laboratories. RT-PCR is also very sensitive so positive results can often be obtained from samples which are negative by culture (Schwarz

et al., 1996). Viral RNA has been detected in serum samples while infective virus is frequently no longer present (Burt et al., 1998). Another advantage of molecular diagnostic assays is that they are rapid in comparison to virus culture, hence a proper diagnosis can be reported within 8 hours of receiving the first specimen (Burt et al., 1998). CCHFV RT-PCR assays have greatly improved epidemiological studies, for example, the ability to detect viral nucleic acid directly from ticks collected in the field. An added benefit of these techniques is that they give way for molecular epidemiology to be executed. Phylogenetic analysis can be performed on amplified viral complementary DNA (cDNA) that has been sequenced. This approach made it possible for the source of a CCHF outbreak in the United Arab Emirates to be determined (Rodriguez et al., 1997).

Conventional RT-PCR assay was further modified by the development of automated real-time assays. The real-time PCR assay has a number of advantages over conventional RT-PCR methods. These include lower rates of contamination, higher sensitivity and specificity, and they are also rapid, making results available in minutes instead of hours. Many researchers have described the use of real-time PCR assays for detecting number of viral causes of haemorrhagic fevers, including Ebola, Rift Valley fever, and dengue viruses. (Drosten et al., 2002).

Figure 6 illustrates the laboratory confirmation of CCHF. During the acute stage of illness virus has been isolated from serum samples collected on days 1 to 12, although levels of viremia are higher during the early acute stage particularly in fatal infections. Viral RNA has been detected up to day 16 in patients with non-fatal infections. Antibody is detectable from day 3 but more frequently from days 5 to 7 in non-fatal cases. Patients with a fatal outcome may not develop detectable antibody responses. It is important to be able to distinguish serologically between an IgG and IgM response to confirm a recent or past infection. ELISA and IFAs are designed to test for a specific IgG or IgM response and therefore can distinguish between them. The presence of IgM antibodies indicates a recent infection and IgG a past infection.

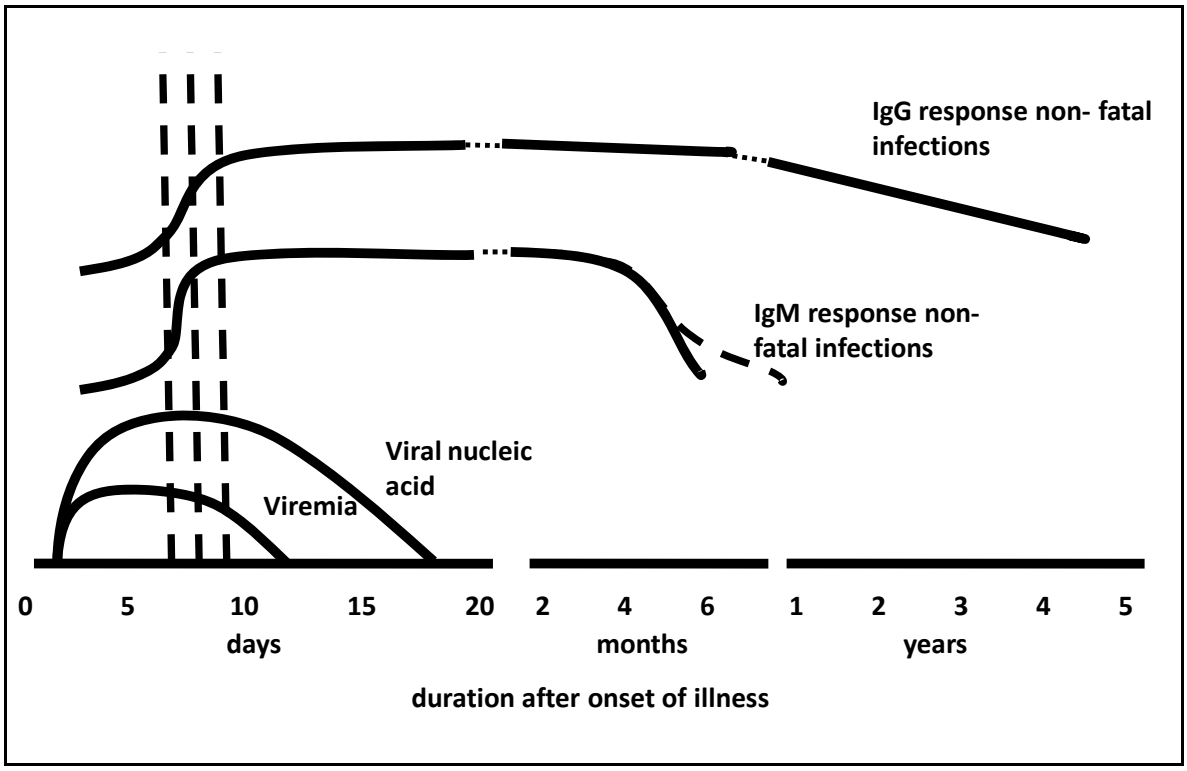


Figure 6. Graph demonstrating laboratory confirmation of CCHFV (Burt, 2011).

1.8 Differential diagnosis

Most of the suspected cases of CCHFV seem to be severe infections with more common agents including bacterial septicaemias, malaria, rickettsias, viral hepatitis and HIV-AIDS related complications. After a clinical diagnosis has been made it is crucial to establish an accurate history of possible exposure to infection, signs and symptoms of illness and clinical pathology results for interpretation of results.

CCHF infections need to be differentiated from other tick-borne pathogens that cause febrile illness in humans (Burt et al., 1996). In Africa, specific consideration needs to be given to tick-borne typhus, commonly known as tick-bite fever caused by *Rickettsia conorri* or *R. africae*. Patients with tick-bite fever often present with characteristic necrotic lesions or eschars at the site of the tick bite. The incubation period is usually 7-10 days which is longer than that for CCHF. Tick-bite fever is associated with a petechial rash and can be fatal in humans with haemorrhagic signs similar to CCHF, but can be treatment with broad-spectrum antibiotics. Other tick-borne diseases to be considered include Q fever caused by *Coxiella burnetti*, ehrlichiosis (*Ehrlichia spp*), borrelia (*Borrelia spp*) or infection with *Babesia* spp.

There are a number of tick-borne viruses in Africa besides CCHF, which have also been associated with human disease, for example Dugbe and Nairobi sheep disease (Burt et al. 1996). In Africa Rift Valley fever virus, a mosquito-borne pathogen can also cause a fatal disease with haemorrhagic manifestations in humans. The disease can be acquired from contact with the tissues of infected livestock and mosquito bite (Gerdes, 2004).

Particular consideration should also be given to the other viral haemorrhagic fevers occurring in Africa. They include Marburg disease and Ebola haemorrhagic fever, caused by members of the *Filoviridae* family, and Lassa fever caused by a virus of *Arenaviridae* family. Marburg and Ebola viruses cause sudden outbreaks of highly fatal disease in tropical Africa usually associated with similar disease in non-human primates, however the natural reservoir of these viruses are still not known. Lassa fever virus is associated with chronic renal infection of *Mastomys* spp. rodents in West Africa, and humans acquire infection through contact with food and house dust contaminated with rodent urine. In September and October 2008, a new member of the *Arenaviridae* family named Lujo virus (LUJV) was isolated in South Africa and responsible for a nosocomial outbreak of disease that killed 4 of 5 patients (Briese et al., 2009; Paweska et al., 2009). LUJV which is highly pathogenic causes haemorrhagic fever and is the first arenavirus from the Old World (OW) to be identified in Africa in the past three decades. Sequencing and phylogenetic analyses however have classified this virus as highly novel and genetically different from the virus species of the OW including the lymphocytic choriomeningitis virus (LCMV) lineage (Briese et al., 2009).

Another group of viruses transmitted by rodents which belong to the genus *Hantavirus* of the *Bunyaviridae* family are endemic in Africa, Asia and the Americas. Hantaviruses of Europe and Asia cause diseases collectively known as haemorrhagic fever with renal syndrome (HFRS) and these viruses could be confused with CCHF from time to time. The hantaviruses from North and South America cause hantavirus pulmonary syndrome (HPS) which less likely to be mistaken for CCHF. There is no conclusive evidence regarding the presence of hantaviruses in Africa, however they are endemic in Europe and Asia and should therefore be included in the differential diagnosis of CCHF.

1.9 Prevention, treatment and control

Hospital acquired infections have been associated with needle stick injuries, or contact of broken skin with infected blood or other tissues and body fluids of infected patients. Transmission via aerosols is not seen as a primary mode of transmission. Patients who are suspected of CCHFV infection should be isolated and subjected to barrier-nursing techniques until the diagnosis is confirmed or excluded. This is necessary to protect health care workers from potential exposure to infection. So the patient has to be isolated in a room with an ante-room or waiting room next door for storing supplies needed for barrier-nursing and care of the patient. Health care workers wear protective clothing such as disposable gowns, gloves, goggles, masks and overshoes which are thrown away on leaving the isolation room via the ante-room. All items that are removed from the isolation ward are disposed of safely or appropriately disinfected. Blood samples are supposed to be wrapped in absorbent material such as paper towels, and kept in secondary leak proof containers or sealed plastic bags for safe transport to the laboratory.

The control of CCHF through the application of pesticides to livestock is not practical. Pyrethroid preparations composed of synthetic chemical compounds are available which can be used to kill ticks which come into contact with human clothing. An outbreak of CCHF occurred in an abattoir in South Africa in 1996, and thereafter it was decided that ostriches should be treated for ticks with pyrethroids and placed in a tick-free environment for two weeks before slaughter to minimize the risk of exposing abbatoir workers to infection (Capua, 1998; Swanepoel et al., 1998). Veterinarians, slaughtermen and other occupations involving livestock should be aware of the disease and take practical steps where applicable, such as wearing gloves, to avoid exposure of naked skin to fresh blood and other animal tissues.

CCHF is treated by means of supportive and replacement therapy with blood products. Immune plasma has been utilized but the efficiency of this treatment is not well established as no systematic investigation has been done with a uniform product of known virus-neutralization activity. Some results showing great potential were obtained in limited trials of ribavirin, a chemotherapeutic drug (Fisher-Hoch et

al., 1995; Ergonul, 2008), but the disease is usually only detected at a late stage and treatment should be started preferably before day five of illness when specific deaths begin to occur.

Ribavirin (Virazole®) is a synthetic purine nucleoside analogue which has a modified base and d-ribose sugar. It was first produced by Sidwell and his colleagues in 1972 (Graci and Cameron, 2006; Hayden, 2006) and inhibits the replication of many RNA and DNA viruses *in vitro*. Ribavirin was the first synthetic nucleoside to show broad spectrum antiviral activity. It is one of the few drugs in clinical use against viruses other than the human immunodeficiency virus and herpesviruses (Graci and Cameron, 2006).

Ribavirin is recommended for the treatment of infected patients with viral haemorrhagic fever syndromes, including CCHF and Lassa fever (McCormick et al., 1986; Ergonul, 2008). Viruses in the *Bunyaviridae* family are usually susceptible to ribavirin even though its mechanism of action is not well understood. An in *vitro* study was performed where ribavirin was shown to inhibit viral activity, and a few of the CCHF viral strains seemed to be more sensitive than others (Watts et al., 1989b). In an experimental study performed on mice, treatment with ribavirin considerably reduced infant mouse mortality and the mean time to death was extended (Tignor and Hanham, 1993). No adverse events related to treatment with ribavirin have yet been documented among CCHF patients due to the acute and short course of the disease, which may not permit time for the development of the side effects, and overshadowing of the signs of CCHFV infection, which are similar as some potential adverse events, for example anemia (Ergonul, 2008).

1. 10 Problem identification, aims and objectives

CCHFV is a serious public health concern with fatality rates in South Africa of approximately 25% and the tendency to cause nosocomial infections. In recent years there has been a significant increase in the number of cases occurring in Balkan countries (Maltezou et al., 2010). Factors such as changes in land use, movement of livestock as well as natural occurrences, for example changes in weather patterns and bird migration can contribute to the emergence and spread of

vector-borne pathogens.

The emergence and re-emergence of CCHFV gives emphasis to the significance of increasing both human and veterinary surveillance and developing diagnostic capacity which needs standardized, rapid and sensitive assays to be developed. Current serological diagnostic assays based on ELISA or IF use inactivated virus which necessitates BSL4 facilities for preparation of reagents. There is a need for a recombinant antigen which is inexpensive to produce and safe to use. Recombinant antigens are safer to use than native antigens and have been shown to still be sensitive and specific in detecting IgM and IgG antibodies against CCHFV (Garcia et al., 2006). Recombinant CCHFV nucleocapsid protein has previously been expressed using baculovirus systems and has been expressed in mammalian cells using a recombinant Semliki Forest alphavirus replicon (Marriot et al., 1994; Saijo et al., 2002a; Tang et al., 2003; Saijo et al., 2005; Garcia et al., 2006).

Marriot et al. (1994), expressed the CCHFV NP of a Greek isolate in a bacPAK6 baculovirus expression system and expressed three NP based peptides in *Escherichia coli* (*E.coli*) as fusions with glutathione *S*- transferase and the antigenic properties of the expressed proteins were tested by ELISA. The baculovirus-expressed protein and the NP based peptides reacted in ELISA with CCHFV antibodies in sera from experimentally infected laboratory animals. The proteins were also able to detect CCHFV specific IgG antibodies when tested with a panel of known positive and negative human sera. A study done by Saijo et al. (2002) described the expression of a recombinant CCHFV NP in a baculovirus expression system. The CCHFV NP was expressed from a Chinese isolate 8402 and detected IgG antibodies in patients previously infected with CCHFV. Mammalian expression of a CCHFV NP protein was described by Garcia et al. (2006). In the study, a recombinant CCHFV NP was expressed in mammalian cells using the recombinant Semliki Forest alphavirus replicon. An indirect IFA and ELISA were developed by immunocapture to detect IgM and IgG antibodies in sera from humans and animals. Analyses of clinical patient samples and animal sera collected in Iran by ELISA and IFA demonstrated the sensitivity and specificity of this antigen for diagnosis of CCHFV infection. However small numbers of serum samples were tested in all these systems without precise dates after onset of illness to indicate usefulness as a diagnostic tool. To the best of our knowledge a recombinant NP antigen has not

been generated using a bacterial expression system for use as a diagnostic tool.

Therefore the aims of our study were to:

- Prepare a recombinant nucleoprotein using a bacterial expression system
- Determine if recombinant protein is functional as a diagnostic tool for detection of IgG antibody responses in survivors of CCHF infection
- Determine if the protein is immunogenic and can induce an antibody response

CHAPTER 2

CLONING AND SEQUENCE ANALYSIS OF THE GENE ENCODING THE NUCLEOPROTEIN OF CCHFV SPU415/85

2.1 Introduction

There are three structural proteins encoded by the CCHFV genome, namely the nucleocapsid, encoded by the S segment and glycoproteins G_n and G_c, encoded by the M segment (Clerx et al. 1981; Elliot et al.1990). CCHFV S RNA comprises approximately 1672 nucleotides and has a single open reading frame which encodes the nucleoprotein (54 x 10^3Da), the major structural protein of the virus (Marriott and Nuttall, 1992). The NP of CCHFV is also believed to be the most antigenic viral protein and hence has been the most commonly used antigen in diagnostic assays.

We selected to clone the ORF of the S segment gene encoding the NP of South African isolate CCHFV SPU 415/85 into pGEM®T Easy vector. The pGEM®T Easy vector is a convenient system for cloning and sequencing PCR products. Genes can subsequently be rescued from these plasmids using restriction enzyme digestion and used for cloning into a suitable expression system using the specified restriction sites on the 5' and 3' ends. The pGEM®T Easy cloning vector has overhanging thymidines added to the 3' ends to facilitate TA cloning of amplicons with overhanging 3' adenines.

2.2 Materials and Methods

2.2.1 Viral RNA

RNA of CCHFV SPU 415/85 was kindly supplied to us by the Special Pathogens Unit (SPU), National Institute for Communicable Diseases of the National Health LaboratoryService (NICD/NHLS) in Johannesburg, South Africa.

2.2.2 Primers

A primer pair was identified to amplify the entire open reading frame (ORF) of the S segment encoding the nucleoprotein. The identified primers were modified to include restriction sites at the 5' ends to facilitate cloning into bacterial expression vectors. The sequences for the restriction sites, namely BamH1 (GGATCC) and Pst1 (CTGCAG), were added to the 5' end of the forward and reverse primer respectively. These restriction sites are highlighted in bold at the 5' end of the oligonucleotide sequences of the primers as shown in Table 1. The primers were selected to contain approximately 40%- 60% G/C content with similar Tm values, which were calculated using the website www.promega.com/biomath. The primers were designated S Bact F (forward) and S Bact R (reverse) with optimal lengths of 37 and 34 bp respectively, including restriction sites.

Table 1. Nucleotide sequences of primers designed for amplification of the CCHFV NP and their positions relative to reference isolate CCHFV SPU 415/85, Accession No. U88415

Primer	Nucleotide sequence	Genomic position	G/C content	Tm
S Bact F	5'GGC **GGATCC** GAAAACAAAATTGAGGTGAATAACAAAG 3'	59-86	42%	46ºC
S Bact R	5' GCC **CTGCAG** GATAATGTTAACACTGGTGGCATTG 3'	1501-1477	50%	48ºC

2.2.3 One step Reverse Transcriptase Polymerase Chain Reaction

The Titan™ One Tube RT-PCR System (Roche, Germany) was utilized for amplifying RNA. It contains AMV reverse transcriptase for cDNA synthesis and Expand High Fidelity enzymes for the amplification process. The reaction mixtures were set up as follows: 0.75µl of a 20 picomolar solution of each primer, 0.25µl (10 U/µl) of protector RNase inhibitor, 2.5µl of 5mM dithiothreitol, 10µl of 5x RT-PCR buffer (7.5mM $MgCl_2$ and dimethyl sulfoxide), 1µl enzyme mix (5U/µl), 2µl template RNA. A negative control reaction was set up in the same way without the addition of template RNA. The reactions were cycled in a Perkin Elmer thermocycler (Gene Amp PCR system 2400, Applied Biosystems) at 50ºC for 30min followed by 30

cycles of denaturation at 94ºC for 10s, annealing at 46ºC for 30s, elongation at 68ºC for 45s, and one final elongation cycle at 68ºC for 7min. The samples were held at 4ºC.

2.2.4 Agarose gel electrophoresis

Electrophoresis of a 10µl aliquot of the PCR products was done using a 1% agarose gel prepared in Tris-Acetate-EDTA (TAE) buffer (pH 8.5) containing ethidium bromide (0.5µg/ml) (Sambrook and Russell, 2001). O'GeneRuler DNA ladder mix comprising DNA fragments from 100 to 10000bp fragments (Fermentas, USA) was used to determine the size of the amplicons. The samples were loaded in a 6x loading dye solution containing 60mM Tris, 10mM EDTA, 0.02% bromophenol blue and 60% glycerol in H_2O. Gel electrophoresis was performed using a BioRad PowerPac Basic system for 1 hour at 90V. The DNA stained with ethidium bromide was visualized with a UV transilluminator.

2.2.5 DNA purification

The Wizard®SV Gel and PCR Clean-Up System (Promega, USA) was used to purify the PCR product directly from the amplicon according to the manufacturer's instructions. This system is based on the ability of the DNA to bind to silica membranes in the presence of chaotropic salts and removes excess nucleotides, primers and enzymes. A 40µl aliquot of the PCR product was added to an equal volume (i.e. 40µl) of membrane binding solution (supplied in the kit). One SV minicolumn was placed in a collection tube. The PCR product mix was transferred to the SV minicolumn assembly and incubated for 1min at room temperature (22-25ºC). The column was centrifuged at 16 000 × g for 1min in a microfuge 16M Spectrafuge, (Labnet International, USA). The column was washed by adding 700µl membrane wash solution (supplied in the kit). The column was centrifuged for 1min at 16 000 × g. The wash was repeated with 500µl membrane wash solution and centrifugation of the column was at 16 000 × g for 5min. An additional spin at 16 000 × g for 5min was done to allow evaporation of any residual ethanol. The SV minicolumn was transferred to a 1.5ml microcentrifuge tube. The DNA was eluted in 50µl of nuclease free water, centrifugation at 16 000 x g for 1 min, quantified and

stored at - 20ºC.

2.2.6 Concentration of DNA

The Quant-iT dsDNA HS Assay Kit with the Qubit fluorometer (Invitrogen, USA) was used to quantitate and measure DNA concentrations according to the manufacturer's instructions. The Quant-iT dsDNA HS Assay is supplied with two standards of the following concentrations: Standard (S)1 = 0ng/µl in 1x TE buffer (10mM Tris, 1 mM EDTA, pH 7.7), S2 = 100ng/µl in TE buffer. A Quant-iT working solution was made by diluting the Quant-iT dsDNA reagent (supplied in the kit) 1:200 in Quant-iT ds DNA HS Buffer (supplied in the kit). An aliquot of 190µl of Quant-iT working solution was loaded into each of the tubes used for preparing the standards. A 10µl aliquot of each Quant-iT standard (supplied in the kit) was added to the appropriate tube and mixed by vortexing for 2-3 sec. A 199µl aliquot of the Quant-iT working solution was added into individual assay tubes and 1µl of DNA sample was added and briefly mixed by vortexing for 2-3 seconds. The final volume in each tube was 200µl. All tubes were allowed to incubate at room temperature for 2 min. The Qubit fluorometer was calibrated for each measure using the two standard solutions prepared. The reading was recorded in ng/ml and the dilution was considered when calculating the final concentration.

2.2.7 Cloning of gene encoding NP into pGEM® T Easy bacterial vector using TA cloning

The high-copy number pGEM®T Easy Vector contains T7 and SP6 RNA polymerase promoters flanking multiple cloning region within the α- peptide coding region of the enzyme ß-galactosidase. The vector map and multiple cloning site (MCS) are shown in Figure 7 and Table 2 shows the sequence reference points of the vector.

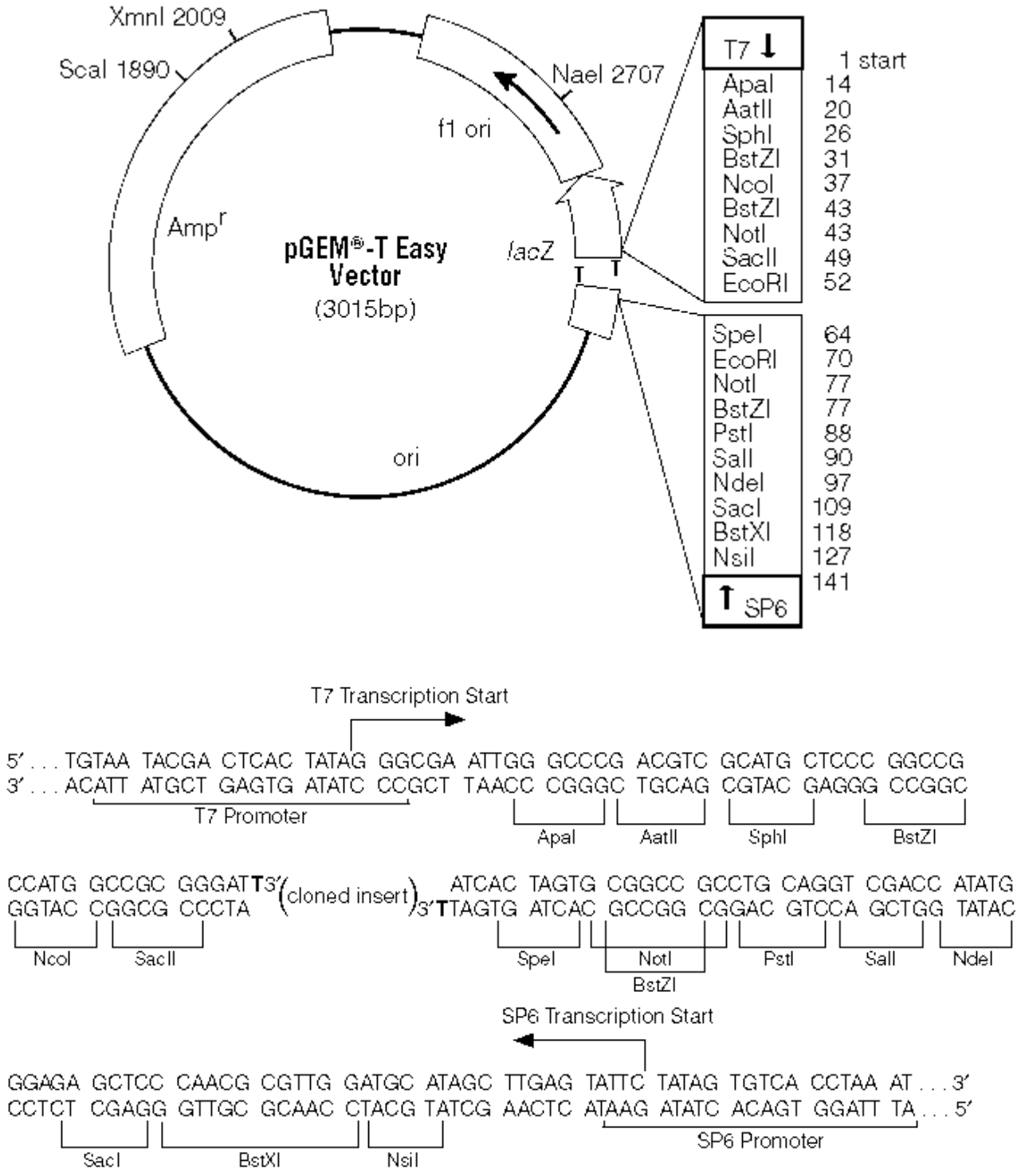


Figure 7. The promoter, multiple cloning site and vector map of the pGEM®T Easy vector.

Table 2. Sequence reference points of pGEM® T Easy vector.

The pGEM®T Easy sequence reference points	Position on vector
T7 RNA polymerase transcription initiation site multiple cloning region	1
Multiple cloning region	10-133
SP6 RNA polymerase promoter (-17 to +3)	124-143
SP6 RNA polymerase transcription initiation site	126
pUC/M13 Reverse sequencing primer binding site	161-177
*lac*Z start codon	165
lac operator	185-201
ß-lactamase coding region	1322-2182
Phage f1 region	2365-2820
lac operon sequences	821-2981,151-380
pUC/M13 forward sequencing primer binding site	2941-2957

The amplicons were ligated into pGEM®T Easy vector by T4 DNA ligase. Chemically competent JM109 *E. coli* host cells (Promega, USA), with a transformation efficiency of 1 × 10^8 cell forming units/µg DNA, were transformed according to the manufacturer's instructions. Blue/ white colony selection was done and the selected transformants were grown overnight (O/N) and purified. Blue/white colony selection is based on the principle that the successful cloning of inserts into the pGEM®T Easy vector will interrupt the action of the *lacZ* gene. ß-galactosidase converts the colourless substrate X-gal (5-bromo-4-chloro-3-indolyl-[beta]-D-galactopyranoside) to produce blue colonies. The *lacZ* gene contains the multiple cloning and A/T cloning sites. The gene will be disrupted in positive transformants, therefore ß-galactosidase will no longer be produced and X-gal can no longer be metabolized to produce blue colonies. Thus colonies containing positive transformants are likely to be white. However it is necessary to confirm positive transformants as cloning may result in blue colonies when the insert is a multiple of 3 bases long (including 3' A overhangs) and does not contain in-frame stop codons thus not disrupting the *lacZ* gene.

2.2.8 Preparation of chemically competent cells

Chemically competent cells were prepared under sterile conditions utilizing the calcium chloride method. Briefly, a 3ml O/N culture of *E. coli* OverExpress C43 (DE3) competent cells (Lucigen, USA) was prepared in Luria Bertani (LB) broth without ampicillin (amp). A 100ml volume of LB broth was prewarmed in a 250ml Erlenmeyer flask and 1ml of the O/N culture was added. The culture was incubated at 37ºC with shaking at 200rpm. The OD_{600} of the *E.coli* culture was determined at 30min intervals starting at time 0. When the OD_{600} reached between 0.45 - 0.5, the cells were transferred to ice. Cells from a 20ml culture volume were collected by centrifugation at 2000xg for 5min at 4ºC. The cells were resuspended in half the original volume (10ml) of freshly prepared ice-cold 50mM $CaCl_2$ (0.11g Ca Cl_2/ 15ml ddH_2O). The cells were then centrifuged for 5min at 2000 × g at 4ºC. The supernatant was poured off and the cells resuspended in 1ml ice-cold $CaCl_2$. The mixture was left on ice for 1 hour, aliquoted into cryotubes in 15% glycerol and frozen at -70ºC. The transformation efficiency of the cells were determined using the formula:

$$\frac{\text{Colony forming units(cfu)}}{\text{ng DNA}} = \text{cfu/ng DNA}$$

2.2.9 Ligation reactions

The CCHFV NP gene was ligated into pGEM®T Easy vector generating recombinant CCHFV NP construct. Ligation reactions prepared are shown in Table 3. A ligation reaction where purified DNA was replaced with deionized water was used as a negative control. The final volume of each ligation reaction was 10µl. The reactions were mixed by pipetting and incubated for 1 hour at 22-25ºC

Table 3. Ligation reaction for cloning CCHFV NP into pGEM®T Easy vector(Promega, USA).

Reaction component	Reaction volume
2x rapid ligation buffer (400mM Tris-HCl, 100mM $MgCl_2$, 100mM DTT, 5mM ATP, pH 7.8)	5µl
T4 DNA ligase (1U/µl)	1µl
Purified CCHFV NP DNA (75ng/µl)	3 µl
pGEM® T Easy plasmid (50ng/µl)	1µl
Total	10µl

2.2.10 Transformation of chemically competent JM109 cells

A 50µl aliquot of chemically competent JM109 cells (1×10^8 cfu/µg DNA)(Promega, USA) were transformed with the plasmid using the heat shock method. This method uses the principle that when cells are made competent, their membranes are modified to facilitate the uptake of the DNA plasmid during the heat shock step. Briefly, 2µl of the ligation reaction was added to a 50µl aliquot of chemically competent cells, gently mixed and placed on ice for 20min. The cells were heat shocked for 45-50 seconds at 42ºC, and immediately transferred to ice for 2 min. A 950µl aliquot of room temperature SOC medium was added to the tube and incubated for 1.5 hours at 37ºC with shaking at 200rpm. SOC medium was prepared as follows: 2g Bacto® tryptone, 0.5g Bacto®- yeast extract, 1ml 1M NaCl and 0.25ml 1M KCl were made up to a total volume of 97ml distilled water. The mixture was autoclaved at 121ºC and cooled to room temperature. A 2M stock solution of $MgCl_2.6H_20$ and $MgSO_4.7H_20$ and 2M stock solution of glucose were prepared and added to the tryptone mixture to give a final concentration of 20mM. Distilled water was added to a final volume of 100ml. The final solution was filter sterilized with a pH of 7.0.

LB/amp plates were prepared as follows: 10g Bacto-tryptone, 5g-Bacto yeast, 10g sodium chloride, 15g agar and distilled water to a final volume of I liter. The broth was autoclaved at 121ºC for approximately 30 minutes and allowed to cool to 50ºC before adding amp to a final concentration of 100µg/ml. Aliquots of 30-35ml of broth

were poured into 85mm petri dishes and allowed to solidify. A 100µl aliquot of 100mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Fermentas, USA) and 40µl of 50mg/ml X-gal (Fermentas, USA) were spread over the surface of an LB amp plate and allowed to adsorb for 30 minutes before use. A 300µl aliquot of cells from the ligation-transformation mix was spread over the plate which was incubated O/N at 37ºC.

2.2.11 Plasmid purification

To identify positive transformants, three white colonies were selected from the amp/IPTG/Xgal plate and each colony, designated colony 1, 2 and 3 respectively, was grown O/N in a 5ml volume of LB/amp at 37ºC with shaking at 200rpm. A blue colony selected and grown in the same way was used as negative control. The following day the plasmids were purified using the QIAprep Spin Miniprep kit (QIAgen USA) according to the manufacturer's instructions. This kit is designed for purification of up to 20µg of plasmid DNA from 5ml overnight cultures. All centrifugation steps were performed at 16000 × g. A 3ml aliquot of O/N culture was pelleted for 1 min. The pelleted bacterial cells were resuspended in 250µl buffer P1 (supplied in the kit) and transferred to a 1.5ml microcentrifuge tube. An aliquot of 250µl of buffer P2 (lysis buffer, supplied in the kit) was added and the tube mixed by inversion 4-6 times. An aliquot of buffer N3 (neutralization buffer, supplied in the kit) was added and immediately mixed by inverting the tube 4-6 times. The mixture was centrifuged for 10 min in a table-top centrifuge. The supernatant was applied to the QIAprep spin column and then centrifuged for 1 min to discard the flow-through. A recommended wash was done on the column to remove endotoxins by adding 500µl buffer PB (supplied in the kit) and centrifuging for 1 min. The flow-through was discarded. The column was then washed with 750µl of Buffer PE (supplied in the kit) and centrifuged for 1 min. After discarding the flow-through, the column was centrifuged for an additional min to remove residual wash buffer. The column was placed in a clean 1.5ml eppendorf tube. To elute the DNA, 50µl of nuclease free water was added to the center of the column, left to stand for 1 min at room temperature and centrifuged for 1 min. The concentrations of the purified DNA were measured as described (MATERIALS AND METHODS, 2.2.6) and the DNA was stored at -20ºC. Glycerol stocks of each recombinant plasmid O/N culture were

prepared (850µl recombinant plasmid O/N culture + 150µl glycerol) and stored in cryotubes (Nunc, Denmark) at -70ºC.

2.2.12 Restriction enzyme digests

To verify insertion of the CCHFV cDNA in the recombinant plasmids, each clone was analyzed by restriction enzyme digestion with restriction enzymes for excision of the CCHFV NP gene. Different restriction enzymes are optimally functional in different buffers with different constituents. Some crucial factors to consider with buffer constituents are pH and concentration of monovalent cation (K^+ or Na^+). When double digestions are performed a compromise may have to be made to select a single buffer compatible for both enzymes to function optimally. Double digestion reactions were set up according to the recommendations on the Promega website (www.promega.com) as shown in Table 4. Reaction mixtures were incubated for 2 hours at 37ºC.

Table 4. Restriction enzyme analysis of CCHFV NP by double digestion (Promega, USA).

Reaction component	Reaction volume
1 × restriction enzyme buffer H (90mM Tris-HCl, 10mM $MgCl_2$, 50mM NaCl, pH 7.8)	2µl
BamH1(10U/µl)	1µl
Pst1(10U/µl)	1µl
Bovine serum albumin (BSA)	1µl
70-80ng plasmid DNA	10µl
Nuclease free water	5µl
Total	20µl

After restriction enzyme digestion the CCHFV gene from colony 3 designated CCHFV NP gene 3 was excised from the agarose gel. The DNA was purified from the gel slice using the Promega Wizard®SV Gel and PCR Clean-Up kit as described in MATERIALS AND METHODS, 2.2.5.

2.2.13 DNA sequencing of CCHFV NP gene in pGEM®T Easy vector

The recombinant plasmid with CCHFV NP gene 3 was sequenced using the T7 and SP6 primer sites available on the vector. DNA sequencing was performed using the Big® Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA). The reaction mixtures were set up according to the recommendations of the kit as described in Table 5.

Table 5. Sequencing reactions prepared according to the recommendation of Big® Dye Terminator Kit (Applied Biosystems, USA)

Reaction component	Reaction volume
Terminator Ready Reaction	2µl
Sequencing primers: forward/reverse (0.8 picomolar/ primer)	4µl
1 × Sequencing buffer	2µl
DNA template (1443bp 58.9 ng)	2µl
Total	10µl

The following thermal cycling conditions were used:

96ºC – 1 min
96ºC – 10 sec
50ºC – 5 sec
60ºC – 4 min
} 30 cycles
4ºC – hold

The reactions were cycled using the Perkin Elmer Thermocycler. EDTA/ethanol precipitation for cleanup of the sequencing reaction was performed. Briefly, the sequencing reaction volumes were adjusted to 20µl with nuclease free water and transferred to 1ml eppendorf tubes that contained 5µl 125mM EDTA and 60µl absolute ethanol. The tubes were vortexed for 5 sec to mix, and allowed to precipitate at 22-25ºC for 15min. The samples were centrifuged at 4ºC for 20 min at 14000 × g. The supernatant was aspirated. A 500µl aliquot of 70% ethanol was added to the tubes which were centrifuged at 4ºC for 10min at 14000 × g. The supernatant was aspirated. The tubes were dried in an incubator at 37ºC O/N with

caps open until the tubes were completely dry. The samples were submitted to The Department of Microbiology, Biochemistry and Food Biotechnology, Faculty of Natural and Agricultural Sciences, University of the Free State (Bloemfontein, South Africa) for gel electrophoresis. The nucleotide sequence data of the complete S segment was edited using ChromasPro Version 1.34. The nucleotide sequence was aligned in ClustalX 1.81 with sequence data retrieved from GenBank and a nucleotide "blast" analysis (www.ncbi.nlm.nih.gov) was performed on the edited sequence using the website www.ncbi.nlm.nih.gov to confirm the identity of CCHFV. Sequencing results were also used to confirm the presence of the BamH1 and Pst1 restriction sites.

2.3 Results

2.3.1 One step RT-PCR for amplification of the gene encoding the NP of CCHFV SPU 415/85

As shown in Figure 8, a 10µl volume of each amplicon was analyzed by gel electrophoresis on a 1% agarose gel. The amplicon in lane 2 is approximately 1500bp according to the molecular weight marker, which is consistent with the expected size of 1443 bp of the NP gene amplicon. There is no visible band in the negative control lane 3 indicating that there was no contamination in the reaction.

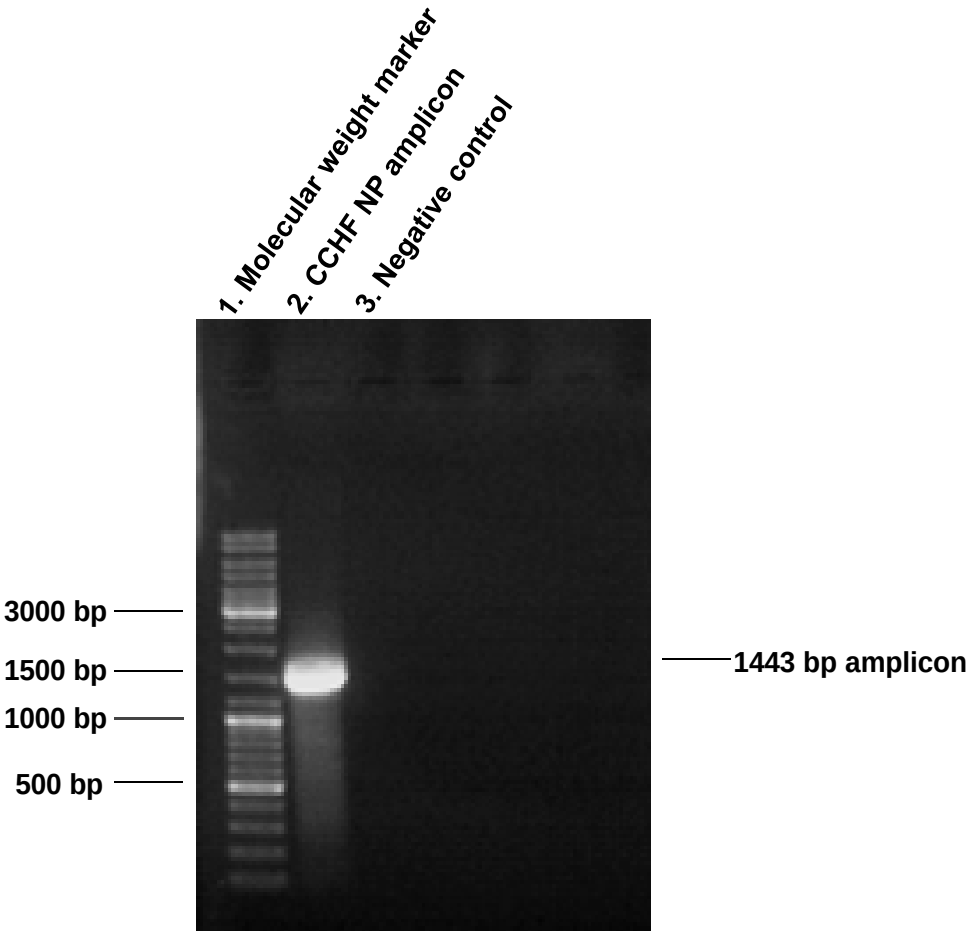


Figure 8. Agarose gel electrophoretic analysis of NP amplicon from the one step RT-PCR, performed on CCHFV viral RNA. Lane 1 – O'GeneRuler DNA ladder mix comprising DNA fragments from 100 to 10000bp fragments (Fermentas, USA) molecular weight marker; Lane 2 - 10µl PCR product of amplified CCHFV RNA using S Bact F and S Bact R primers; Lane 3 – Negative control containing no viral RNA.

After purification, a 5µl volume of the purified product was analyzed by gel electrophoresis on a 1% agarose gel as seen in Figure 9. A strong band of predicted size was present in lane 2. The DNA concentration determined using the Quant-iT dsDNA HS Assay Kit with the Qubit fluorometer was 152.5 ng/µl.

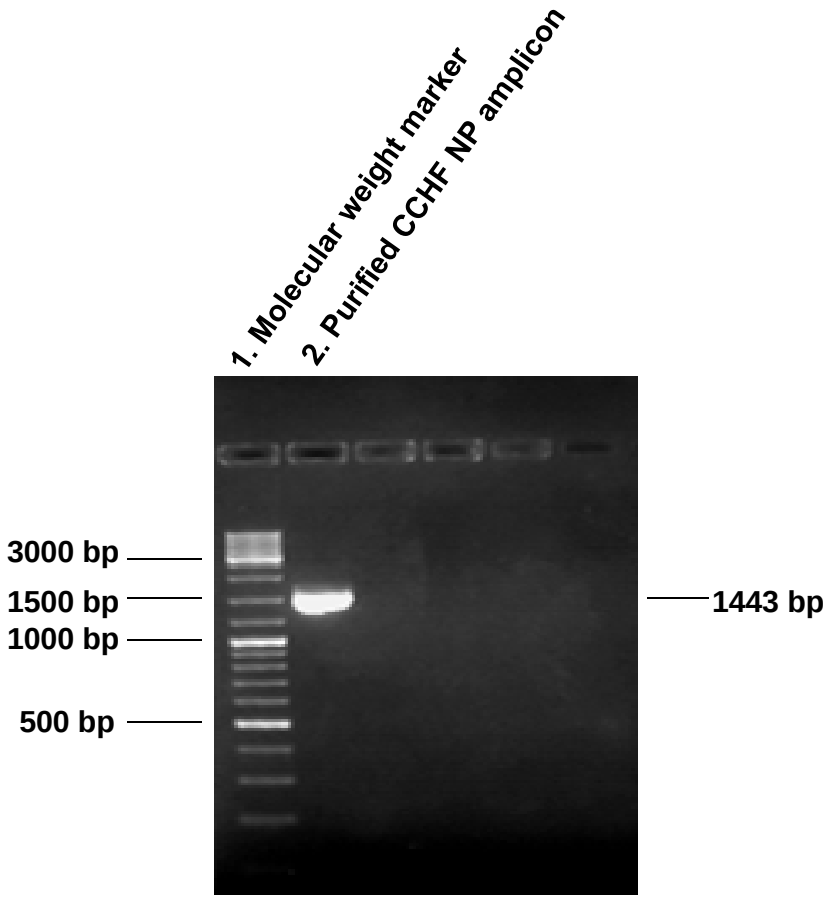


Figure 9. Agarose gel electrophoretic analysis of NP amplicon from the one step RT-PCR after Promega Wizard® SV Gel and PCR Clean-Up. Lane 1 – O'GeneRuler DNA ladder mix molecular weight marker; Lane 2 – 5µl of purified PCR product

2.3.2 A/T cloning of the CCHFV SPU415/85 NP amplicon into pGEM® T Easy vector

Three colonies designated colony 1, colony 2 and colony 3 were selected from the transformation reaction described in MATERIALS AND METHODS, 2.2.10. Results obtained from restriction enzyme analysis of each plasmid are shown in Figure 10. According to the migration patterns of plasmids, as expected the undigested plasmids migrated further (lanes 2, 4 and 6) than the linearied plasmids (lanes 3, 5 and 7) but were identical to the negative control in lane 8 at 3000bp. In the lanes with undigested plasmid bands present at approximately 6000bp likely represent nicked open circular plasmid DNA or relaxed circular plasmid DNA.

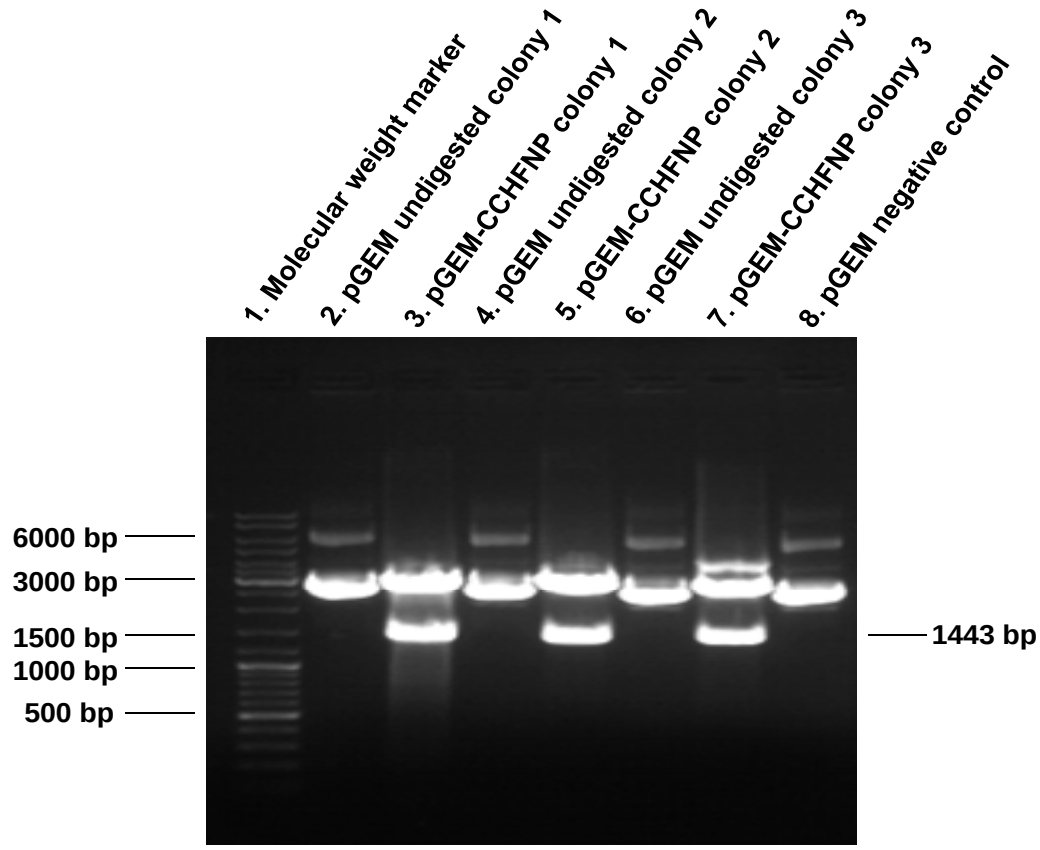


Figure 10. Agarose electrophoretic gel of restriction enzyme analysis of plasmids obtained from the ligation of the CCHFV NP gene into pGEM®T Easy vector.

Lane 1 – O'GeneRuler DNA ladder mix molecular weight marker; Lane 2 – pGEM undigested colony 1; Lane 3 – pGEM-CCHFV NP colony 1 digested BamH1 and Pst1, Lane 4 - pGEM undigested colony 2; Lane 5 - pGEM-CCHFV NP colony 2 digested BamH1 and Pst1; Lane 6 - pGEM undigested colony 3; lane 7 - pGEM-CCHFV NP colony 3 digested BamH1 and Pst1; Lane 7 - pGEM negative control.

Figure 10 shows that double digestion of the recombinant CCHFV NP plasmids with BamH1 and Pst1 yielded the predicted 1443 bp fragments for all three colonies. This confirmed that each construct selected contained the CCHFV NP gene. This was confirmed with DNA sequencing. After purification of the excised band for CCHFV gene 3, a 5µl aliquot of the CCHFV gene 3 purification was analyzed by gel electrophoresis on a 1% gel as shown in Figure 11. The band present in lane 2 of 1443 bp confirms the presence of a band of correct expected size (1443bp). The DNA concentration determined (MATERIALS AND METHODS, 2.2.6) was 74.6ng/µl.

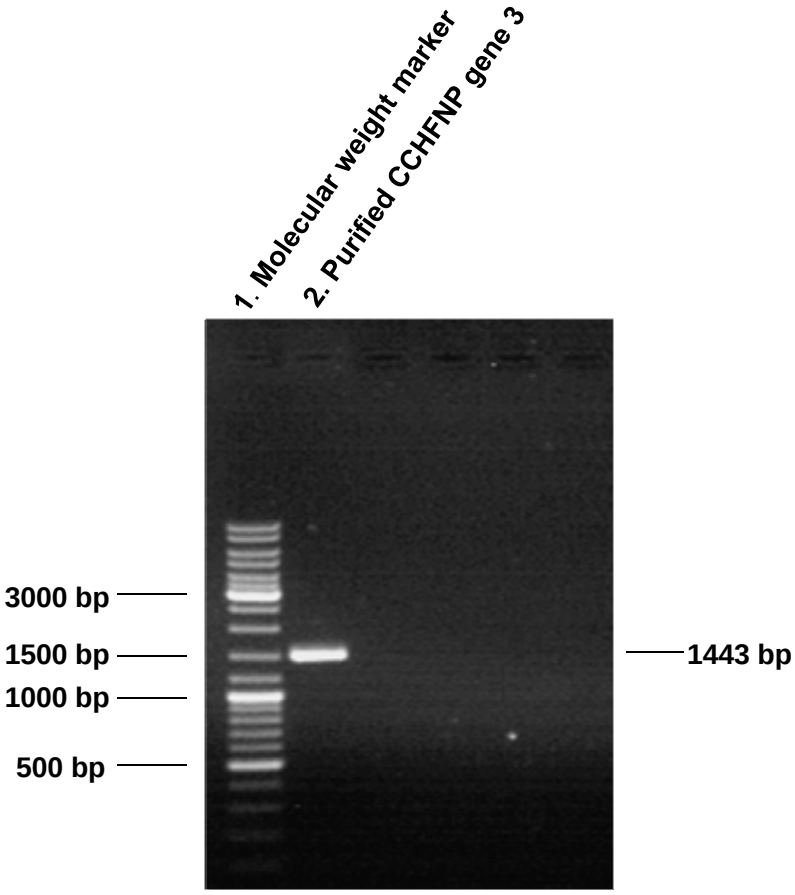


Figure 11. Agarose gel electrophoretic analysis of NP amplicon from the one step RT-PCR after Promega Wizard®SV Gel and PCR Clean-Up. Lane 1 – O'GeneRuler DNA ladder mix molecular weight marker; Lane 2 – 5, 5µl of 30µl clean-up reaction.

2.3.3 Sequencing and sequence analysis of the gene encoding CCHFV SPU415/85 in pGEM®T Easy vector

The nucleotide sequence of CCHFV NP gene 3 construct in the recombinant plasmid is shown in Figure 12. The primers used in the RT-PCR were designed to exclude the start and stop codons on the ORF as they are already present on the pQE-80L bacterial expression vector discussed in the following chapter. A base difference was identified in CCHFV NP gene 3 (Figure 12) compared with sequence data available on GenBank whereby a cysteine (C) was substituted with a thymine (T) at base 1365 of the gene resulting in a change in the codon from CAC to CAT. Nevertheless this was a silent mutation as it did not alter the amino acid of the gene. The nucleotide blast performed on the edited sequence confirmed the identity of the gene as CCHFV. Analysis of the CCHFV NP gene in Nebcutter

(www.nebcutter.com) showed no BamH1 and Pst1 restriction sites within the gene and confirmed the presence of these two restriction sites at the 5' and 3' end of the gene which will be used to facilitate rescue of the gene from the plasmid. Therefore sequence analysis verified insertional cloning of the correct DNA fragment of 1443bp.

Figure 12. Nucleotide sequence of CCHFV NP gene 3 in pGEM. The forward and reverse primers are highlighted in **yellow** with BamH1 and Pst1 restriction sites higlighted in **red** on the forward and reverse primer respectively. The T7 and SP6 transcription sites are highlighted in **pink**.

```
                                1                                                60
CCHFV NP3 SPU415/85 (pGEM)      GATCCCGGGGCCGGCCATGGGCGGCCGCGGGAATTCGATTGCCGGATCCGAAAACAAAAT

                                61                                              120
                                TGAGGTGAATAACAAAGATGAGATGAACAAGTGGTTTGAAGAGTTCAAAAAAGGAAATGG

                                121                                             180
                                ACTTGTGGACACCTTCACAAACTCCTATTCCTTTTGTGAGAGTGTTCCAAATTTGGACAA

                                181                                             240
                                GTTTGTGTTCCAAATGGCCAGTGCCACCGATGATGCACAAAAGGATTCCATCTACGCGTC

                                241                                             300
                                TGCTCTGGTGGAGGCAACAAAATTTTGTGCACCTGTATATGAGTGTGCATGGGTTAGCTC

                                301                                             360
                                CACTGGCATTGTGAAAAGGGGACTTGAATGGTTCGAAAAAAATGCGGGCACCATTAAGTC

                                361                                             420
                                TTGGGATGAAAGTTATACTGAGCTAAAAGTCGACGTCCCAAAAATAGAACAGCTAGCCAA

                                421                                             480
                                TTACCAACAGGCTGCCTTGAAATGGAGGAAAGACATAGGTTTCCGTGTCAATGCAAACAC

                                481                                             540
                                AGCGGCTCTGAGTAACAAAGTCCTCGCAGAGTACAAAGTTCCTGGAGAGATTGTGATGTC

                                541                                             600
                                TGTCAAAGAGATGCTGTCAGACATGATTAGGAGAAGGAACCAGATTCTAAACAGGGGTGG

                                601                                             660
                                TGATGAGAATCCACGTGGCCCTGTGAGCCGTGAGCATGTAGACTGGTGCAGGGAGTTTGT

                                661                                             720
                                CAAAGGCAAATACATCATGGCCTTCAACCCACCATGGGGGGACATCAACAAGTCAGGCCG
                                721                                             780
```

TTCAGGAATAGCACTTGTCGCAACAGGCCTTGCCAAGCTTGCAGAGACTGAGGGAAAGGG

781 840
AGTATTTGACGAAGCCAAAAAGACTGTGGAAGCCCTCAATGGGTATCTGGACAAGCACAA

841 900
GGACGAAGTTGACAGAGCGAGTGCTGACAGCATGATAACAAACCTTCTCAAGCACATTGC

901 960
TAAGGCACAGGAGCTTTATAAGAATTCGTCTGCACTCCGTGCACAAGGTGCACAGATTGA

961 1020
CACTGCTTTCAGCTCATACTATTGGCTTTACAAGGCTGGCGTGACCCCAGAGACATTCCC

1021 1080
GACGGTGTCACAGTTCCTCTTTGAGCTAGGAAAACAGCCAAGAGGTACCAAGAAAATGAA

1081 1140
GAAGGCTCTGCTGAGCACCCCAATGAAGTGGGGGAAGAAACTTTATGAACTCTTTGCCGA

1141 1200
CGATACTTTCCAGCAGAACAGGATCTACATGCACCCTGCCGTGCTTACAGCTGGCAGAAT

1201 1260
CAGTGAAATGGGAGTCTGCTTTGGGACAATCCCCGTGGCCAACCCTGATGATGCTGCTCA

1261 1320
AGGGTCTGGACATACCAAGTCTATTCTCAACCTCCGGACTAACACCGAGACCAACAATCC

1321 1380
GTGTGCCAGGACCATTGTCAAGCTGTTTGAAATTCAGAAAACAGGGTTCAACATTCAGGA

1381 1440
CATGGACATAGTGGCCTCTGAGCACTTGCTGCATCAGTCTCTTGTTGGCAAGCAATCTCC

1441 1500
ATTCCAGAACGCCTACAACGTCAAGGG**CAATGCCACCAGTGTTAACATTATCCTGCAGGG**

1501 1560
CAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCG

2.4 Summary

The S segment of CCHFV encodes the nucleoprotein, the most abundant protein inducing a high immune response (Garcia et al., 2006). The complete open reading frame for the NP gene, excluding the start and stop codons, was amplified from viral RNA isolated from a South African CCHFV isolate (SPU 415/85). The amplicon was subsequently cloned into vector pGEM®T Easy using T/A cloning and the complete nucleotide sequence of the gene determined. Nucleotide sequence data was used to confirm that the amplified gene had no PCR errors that would result in non-synonomous amino acid changes when translated and that the restriction sites engineered on the primers were present The resulting plasmid was stored for downstreaming cloning and expression studies. In addition the sequence analysis provided complete S segment nucleotide sequence data for synthesis of a gene codon optimized for expression in a bacterial host system.

CHAPTER 3

BACTERIAL EXPRESSION OF RECOMBINANT CCHFV NP FROM NATIVE AND CODON OPTIMIZED GENES

3.1 Introduction

The most commonly utilized expression systems are based on bacterial, baculovirus or mammalian expression. The choice of expression system is determined by the specific applications for which the protein is being made. For maximal expression yield bacterial expression is recommended provided that post-translational modifications of the recombinant protein such as phosphorylation, glycosylation, precursor processing, and targeting are not required for protein function. Baculovirus vectors used for heterologous protein expression in insect cells have also become widely utilized (Miller, 1988; Luckow 1991). The principle of this system is based on the ability of the baculovirus to infect and multiply in cultured insect cells. The target gene is cloned into a plasmid transfer vector and then co-transfected with double-stranded baculovirus DNA into insect cells. Many post-translational modifications typically encountered in mammals cells mentioned above also occur in insect cells. However the preparation of baculovirus is relatively expensive and technically more difficult than bacterial expression.

Despite the very low expression levels, mammalian cells are often the best host for the expression of recombinant vertebrate proteins as they produce the same post-translational modifications and recognize the same signals for synthesis, processing, and secretion utilized in the organism from which the sequence was originally derived. In general, they contain an efficient promoter element for high-level transcription initiation, mRNA processing signals such as mRNA cleavage and polyadenylation sequences, selectable markers to select mammalian cells that have stably integrated the DNA into their genome, and plasmid sequences that facilitate the propagation of the vectors in bacterial hosts. However, the yield is seldom appropriate for downstream applications such as development of an ELISA for use as diagnostic surveillance tools where high yields are essential.

E. coli is one of the most widely used prokaryotic hosts for production of foreign proteins (Baneyx, 1999; Gustafsson et al., 2004; Jana and Deb, 2005; Sadhev et al., 2008). This is because high-level expression of protein can be achieved at low cost and because an increasingly large number of vectors and mutant host strains are readily available (Baneyx, 1999). However the production of biologically active proteins in *E. coli* is quite challenging due to the lack of post translational modifications (Jana and Deb, 2005).

A well designed recombinant expression system contains essential genetic elements such as the origin of replication (ori), an antibiotic resistance gene, transcriptional promoters, translation initiation regions as well as transcriptional and translational terminators (Sorensen and Mortensen, 2005). These elements all contribute to expression of foreign proteins in prokaryotic cells.

One of the limitations of prokaryotic expression systems is that the codon usage in *E. coli* displays a bias, indicating a non-random usage of synonomous codons (Kane, 1995; Rocha, 2004; Sharp et al., 2005). This is because a small subset of codons are used that are recognized by the most abundant tRNA species. The force that modulates this codon bias is called translational selection and its strength is important in fast-growing bacteria. As a result heterologous genes that contain codons rarely used in *E. coli* may be inefficiently expressed (Lithwick and Margalit, 2003; Rocha, 2004).

Codon usage optimization basically involves the alteration of the rare codons in the gene to be expressed so that they reflect the codon usage of the host without changing the amino acid sequence of the encoded protein (Gustafsson et al., 2004). The information used for the optimization process is usually the DNA or protein sequence to be optimized and a codon usage table or reference set of the host (Puigbo et al., 2007).

Codon bias can be evaluated by various indices, the Codon Adaptation Index (CAI) (Sharp and Li, 1987) and effective number of codons (Enc) (Wright, 1990). CAI measures the similarity between the codon usage of a gene and the codon usage of a reference group of genes (Sharp and Li, 1987). Its values range from 0 (when

the codon usage of a sequence and that of the reference set are very different) to 1 (when both codon usages are the same). This index is the most effective and most commonly used of all codon bias measures for predicting gene expression levels (Lithwick and Margalit, 2003; Henry and Sharp, 2007) and is the index that was used in this study. This chapter describes the attempts to express the CCHFV NP in a bacterial system using a native gene and a codon optimized gene.

3.2 Materials and methods

3.2.1 Cloning of CCHFV NP gene into pQE-80L bacterial expression vector using restriction sites

The expression system pQE-80L (QIAgen, USA) which expresses proteins at 37 °C was chosen as the vector for expression of CCHFV NP. This expression system utilizes high-level expression of 6xHis-tagged proteins in *E. coli* based on the T5 promoter transcription–translation system. Unlike the pGEM®T Easy Vector, pQE-80L does not utilize blue/white colony selection for identification of positive transformants. All colonies that grow on transformation plates are white.

The pQE-80L vector has the following features which are shown in Figure 13:

- Optimized promoter–operator element consisting of phage T5 promoter (recognized by the *E. coli* RNA polymerase) and two *lac* operator sequences which increase *lac* repressor binding and ensure efficient repression of the powerful T5 promoter
- Synthetic ribosomal binding site, RBSII, for high translation rates
- A 6xHis-tag coding sequence 5' to the cloning region
- Multiple cloning site and translational stop codons in all reading frames for convenient preparation of expression constructs
-Two strong transcriptional terminators: *t0* from phage lambda and T1 from the *rrnB* operon of *E. coli* to ensure stability of the expression construct
- β-lactamase gene (*bla*) which confers resistance to ampicillin at 100µg/ml
- ColE1 origin of replication

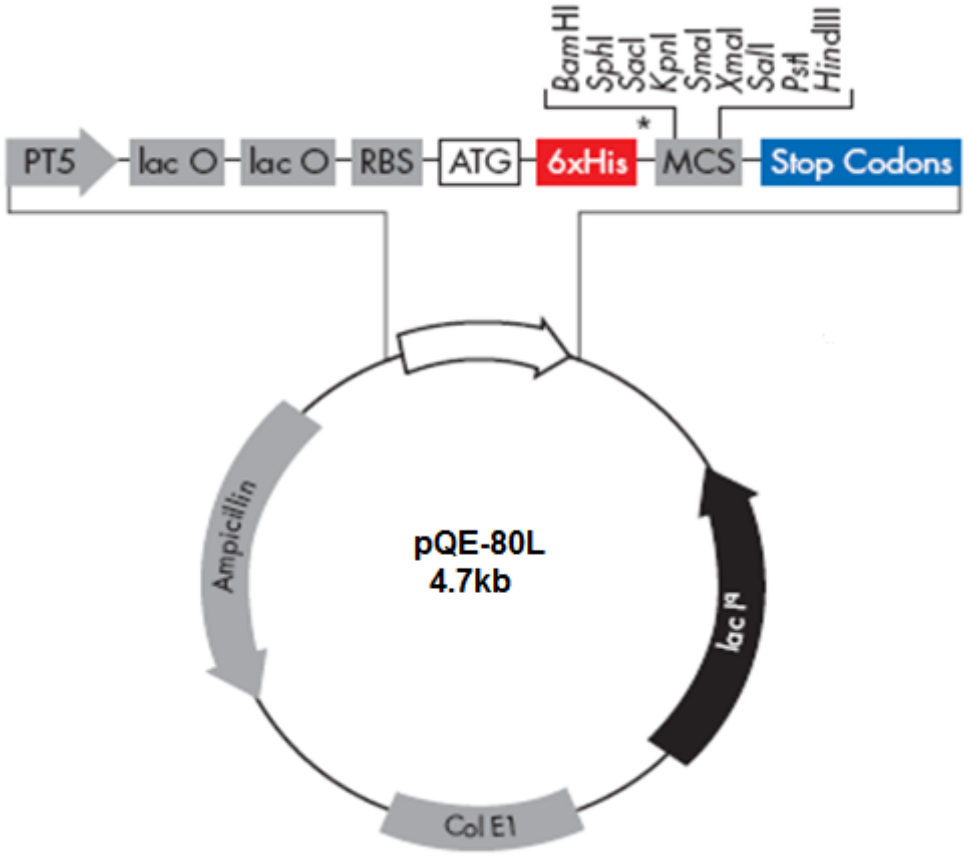


Figure 13. Vector map of pQE-80L bacterial expression system (QIAGEN, USA)

The pQE-80L plasmid DNA was kindly supplied to us by Dr Knox from Rhodes University, Grahamstown, South Africa. Restriction enzyme digestion was performed on the pQE-80L plasmid using BamH1 and Pst1 restriction sites to facilitate ligation of the CCHFV NP gene 3 into pQE-80L.

Purified CCHFV NP gene 3 was rescued from the recombinant plasmid using the restriction sites (BamH1 and Pst1) engineered on the 5' and 3' ends of the amplicon. A ligation reaction was performed to insert CCHFV NP gene 3 into pQE-80L bacterial expression vector using compatible restriction sites (BamH1 and Pst1) and set up according to the recommendations on the Fermentas website (www.fermentas.com) as shown in Table 6. The ligation reaction was mixed by pipetting and incubated for 1 hour at 22-25ºC. The ligation mixture (10µl) was used to transform chemically competent JM109 cells (1 x 10^8 cfu/µg DNA)(Promega, USA) using the heat shock method described in MATERIALS AND METHODS (2.2.10). A 1µl aliquot of pQE-80L plasmid DNA with no gene insert was used to

transform cells as a control. Colonies were selected from respective agar plates containing the appropriate antibiotics to identify positive transformants (MATERIALS AND METHODS, 2.2.11). The recombinant CCHFV plasmids were propagated, purified (MATERIALS AND METHODS, 2.2.11) and a selected positive transformant designated pQE-80L-CCHFV NP was confirmed using gel electrophoresis and restriction enzyme site analysis with BamH1 and Pst1 restriction enzymes (MATERIALS AND METHODS, 2.2.12). A glycerol stock of the positive transformant was also prepared as previously described (MATERIALS AND METHODS, 2.2.11). The pQE-80L control plasmid was prepared similarly.

Table 6. Ligation reaction for cloning CCHFV NP into pQE-80L (Fermentas, USA).

Reaction component	Reaction volume
1 x T4 ligase buffer (400mM Tris-HCL, 100mM $MgCl_2$, 100mM DTT, 5mM ATP, pH 7.8)	1µl
T4 ligase enzyme (1U/µl)	1µl
Purified double digested CCHFV NP gene 3 DNA (65.2ng)	6 µl
Purified double digested pQE-80L plasmid (65.5ng)	2µl
Total	10µl

3.2.2 DNA sequencing of CCHFV NP gene in pQE-80L vector

DNA sequencing of the CCHFV NP gene in pQE-80L plasmid was performed using the Big®Dye Terminator v3.1 Cycle Sequencing Kit as described in MATERIALS AND METHODS, (2.2.13) with a pair of primers that target regions of the pQE-80L plasmid as shown in Table 7 (QIAGEN, USA).

Table 7. Sequencing primers for pQE-80L bacterial expression vector (QIAGEN, USA).

Sequencing primer name	Primer sequence	Genomic position
pQE-fwd (forward)	5’ CCC GAA AAG TGC CAC CTG 3’	1728 – 1745
pQE-rev (reverse)	5’ GTT CTG AGG TCA TTA CTG G 3’	3882 – 3864

3.2.3 Bacterial expression of pQE-80L-CCHFV NP using IPTG induction

An O/N culture of the recombinant pQE-80L-CCHFV NP was prepared from the glycerol stock (MATERIALS AND METHODS, 3.2.1). A 5ml volume of LB/amp was inoculated with 10µl of the glycerol stock of pQE-80L-CCHFV NP. Another 5ml volume of LB/amp was inoculated with 10µl glycerol stock of pQE-80L with no gene insert and used as the negative control. The following day the cultures were diluted 1:20 with pre-warmed LB/amp (2ml culture in 38ml LB/amp) and incubated for approximately 2 hours. The optical density (OD) of the cells were measured with a spectrophotometer (Amersham Biosciences, England). When the OD_{600} reached between 0.6 – 0.8, the cultures were induced with IPTG to a final concentration of 1mM and incubated at 37ºC with shaking at 200rpm. A volume of 1ml of the induced cultures was collected at t=0, t=1h, t=3h and t=24h post-induction. The cells were harvested from the collections by centrifugation at 16 000 × g for 10min using the bench top centrifuge. The supernatant was discarded and the cells resuspended. The volume of PBS (pH 7.4) required for resuspension was calculated using the formula, Volume of PBS = $A_{600nm}/0.5 \times 150\mu l$. In attempts to express the CCHFV NP in pQE-80L vector a number of conditions regarding temperature and concentrations of IPTG were optimized. These conditions are shown in Table 8.

Table 8. Optimization of temperature and IPTG concentrations for expression of CCHFV NP gene in pQE-80L bacterial expression vector.

Final IPTG conc.	Temp. of culture O/N incubation	Temp. of culture post induction
0.01mM	25ºC	25ºC
0.01mM	37ºC	37ºC
0.05mM	25ºC	37ºC
1mM	25ºC	25ºC
2mM	37ºC	25ºC
3.5mM	25ºC	25ºC
5mM	37ºC	37ºC

3.2.4 SDS-Polyacrylamide gel electrophoresis

To separate the proteins, sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) of the samples was performed as described by Laemmli (1970). The resolving gel was prepared as shown in Table 9. The resolving gel was poured into the pre-assembled electrophoresis apparatus (BIORAD, USA). A layer of isobutanol was applied to the top before polymerization for approximately 15min. After the resolving gel solidified the layer of isobutanol was discarded.

Table 9. Preparation of 8% resolving gel for analysis of proteins by SDS-PAGE.

Reaction component	Reaction volume
30% acrylamide solution/0.8% bisacrylamide stock solution (Merck, USA)	21.4ml
1M Tris-HCl pH 8.8 stock solution	30ml
10% SDS	0.8ml
1.5% ammonium persulfate (freshly prepared)	4ml
TEMED	0.02ml
Distilled H_2O	Up to final volume
Total	80ml

The stacking gel was prepared as shown in Table 10. The stacking gel was applied to the top of the resolving gel and the comb inserted before polymerization for approximately 10min.

The tanks of the electrophoresis apparatus were filled with 1× Tris-Glycine-SDS electrophoresis buffer (25mM Tris, 192mM glycine, 0.1% SDS, pH 8.3) (Fermentas, USA). The protein samples were prepared as follows: 20µl protein sample, 8µl of 5x protein loading buffer (0.313 M Tris-HCl, pH 6.8, 10% SDS, 0.05% Bromophenol blue, 50% glycerol) and 1µl 2M DTT (Fermentas, USA). The samples were heated for 5 minutes at 95ºC and 25µl aliquots of the samples were loaded into each well. The gel was run at 180V for 50 minutes. Prestained protein marker (Fermentas, USA) comprising proteins ranging in size from 10 to 130kDa proteins was used to determine the sizes of the expressed proteins.

Table 10. Preparation of 4% stacking gel for analysis of proteins by SDS-PAGE.

Reaction component	Reaction volume
30% acrylamide solution/0.8% bisacrylamide stock solution	2ml
1M Tris-HCl pH 6.8 stock solution	1.9ml
10% SDS	0.15ml
Glycerol	1ml
1.5% ammonium persulphate	0.7ml
TEMED	0.02ml
Distilled H_20	Up to final volume
Total	15ml

SDS-PAGE gels were stained O/N with gentle shaking in approximately 200ml of a staining solution containing 45% methanol, 10% glacial acetic acid, 0.2% Coomasie Brilliant Blue. The stain was discarded and the gels were destained with approximately 600ml destaining solution containing 45% methanol and 7% glacial acetic acid. The gels were then dried in a gel dryer for approximately 2 hours at 70ºC.

3.2.5 Western blot analysis of His tagged CCHFV NP

For Western blot analysis the BM Chemiluminescence Western blotting Kit (Mouse/Rabbit) (Roche, Germany) was used The gels were soaked in transfer buffer (25mM Tris, 192mM glycine, 20% methanol, pH 7.4) for 15min. PVDF nitrocellulose membrane was moistened with methanol for 1-3 seconds, incubated in distilled water for 1-2 minutes and then soaked in transfer buffer for 2-3 minutes. The membrane was transferred to a semi-dry blotter (BIORAD, USA) and run at 15V for 15 min for transfer of proteins from the gel to the membrane. The membrane was removed from the transfer apparatus and washed twice with Tris buffered saline (TBS) buffer (50mM Tris, 150mM NaCl, pH 7.5) for 10min to remove the methanol. The membrane was blocked with 12.5ml of 1% blocking reagent (supplied in the kit) diluted in TBS buffer for 60min at 22-25ºC. The membrane was incubated in primary antibody solution, mouse anti-His_6 IgG (100µg/ml) diluted 1:200 in blocking reagent for 60min with shaking at 22-25ºC. The membrane was

washed twice with TBST (TBS+ 0.1% Tween 20) for 20 min each. The membrane was incubated in secondary antibody solution, rabbit anti-mouse IgG (50U/100µl) diluted 1:1250 in blocking reagent for 30 min with shaking at 22-25ºC. The membrane was washed four times with TBST for 15min each. A 60µl aliquot of starting solution (supplied in the kit) was added to 6ml of substrate (supplied in the kit) to make up the detection reagent which was added to the membrane for 60 sec. The membrane was placed in a clear plastic wrap and exposed to film (Thermo Scientific, USA) for approximately 60 seconds. The film was incubated for 2 minutes in developer (Agfa, Germany) and for 1 minute in fixer (Agfa, Germany). The film was washed with water to remove excess fixer and visualized.

3.2.6 Codon optimization of CCHFV NP gene

The complete nucleotide sequence for the open reading frame encoding the NP gene for CCHFV isolate SPU 415/85 as determined in Chapter 2 was analysed using the Rare Codon Analysis Tool software available on the GenScript website (http://www.genscript.com). The software optimizes codon usage and GC content and eliminates polyadenylation sites, splicing sites, killer motifs and RNA secondary structure. The sequence data was submitted to GenScript for synthesis of a codon optimized gene.

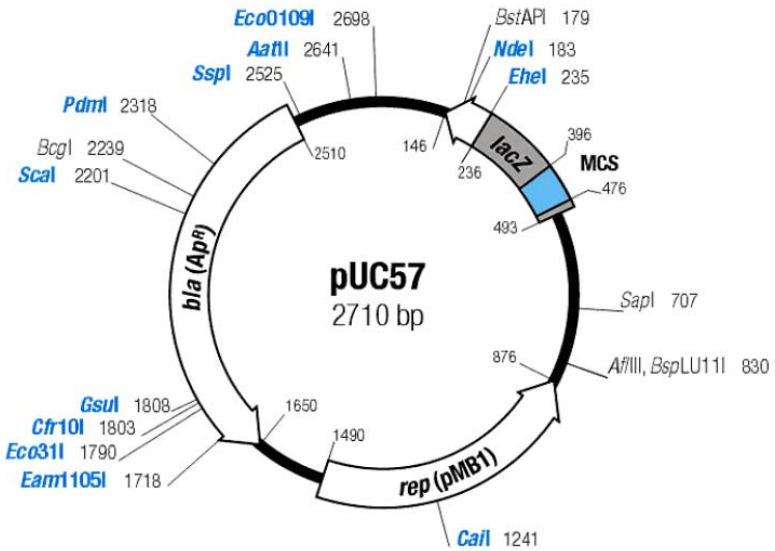
The optimized gene was synthesized by GenScript and supplied cloned in the MCS of pUC57 with BamH1 and Pst1 restriction site modifications at the 5' end and 3' end respectively. Determination of the CAI of the native and optimized genes were made using the GenScript software. The vector map is shown in Figure 14.

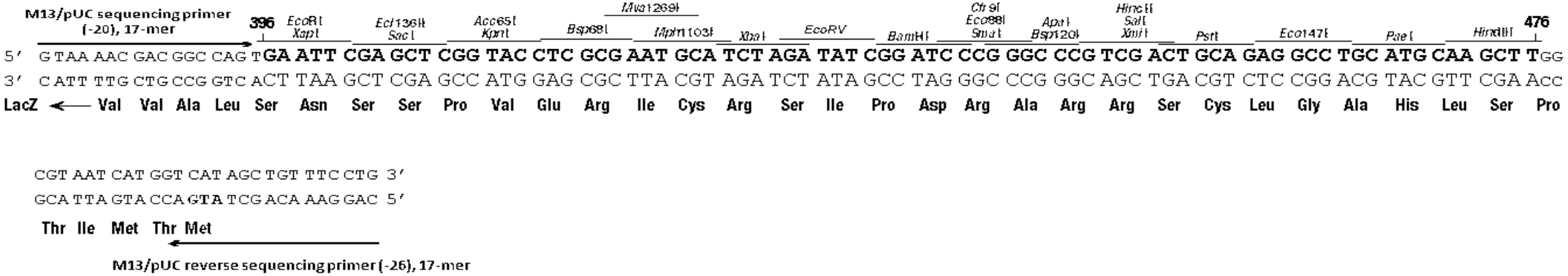
Restriction enzyme digestion was performed on the pCold TF plasmid (Takara Bio, USA) using BamH1 and Pst1 restriction sites to prepare the plasmid for ligation reactions with the codon optimized CCHFV NP gene.

3.2.7 Cloning of gene encoding codon optimized CCHFV NP from pUC57 into pCold TF bacterial expression vector

Takara's pCold TF DNA Vector (Takara Bio, USA) is a fusion cold shock expression vector that expresses Trigger Factor (TF) chaperone as a soluble tag. Trigger Factor is a prokaryotic ribosome-associated chaperone protein (approximately 48kDa) which facilitates co-translational folding of newly expressed polypeptides. TF is highly expressed in *E. coli* expression systems due to its *E. coli* origin. This vector does not make use of blue/white colony selection therefore all colonies growing on plates are white. The pCold TF DNA Vector comprises of a *csp*A promoter plus additional downstream sequences including a 5' untranslated region (5' UTR), a translation enhancing element, a 5' His-tag sequence, and an MCS. A lac operator is inserted downstream of the *csp*A promoter to ensure strict regulation of protein expression. There are also recognition sites for HRV 3C Protease, Thrombin, and Factor Xa located between TF Tag and the MCS which function to facilitate removal of the tag from the expressed fusion protein. The pCold TF vector is optimal for expression of toxic or transmembrane proteins using the "cold shock" technology. The vector map and MCS of pCold TF is shown in Figure 15.

Figure 14. Vector map and multiple cloning sites of cloning vector pUC57 (GenScript Corporation, USA).





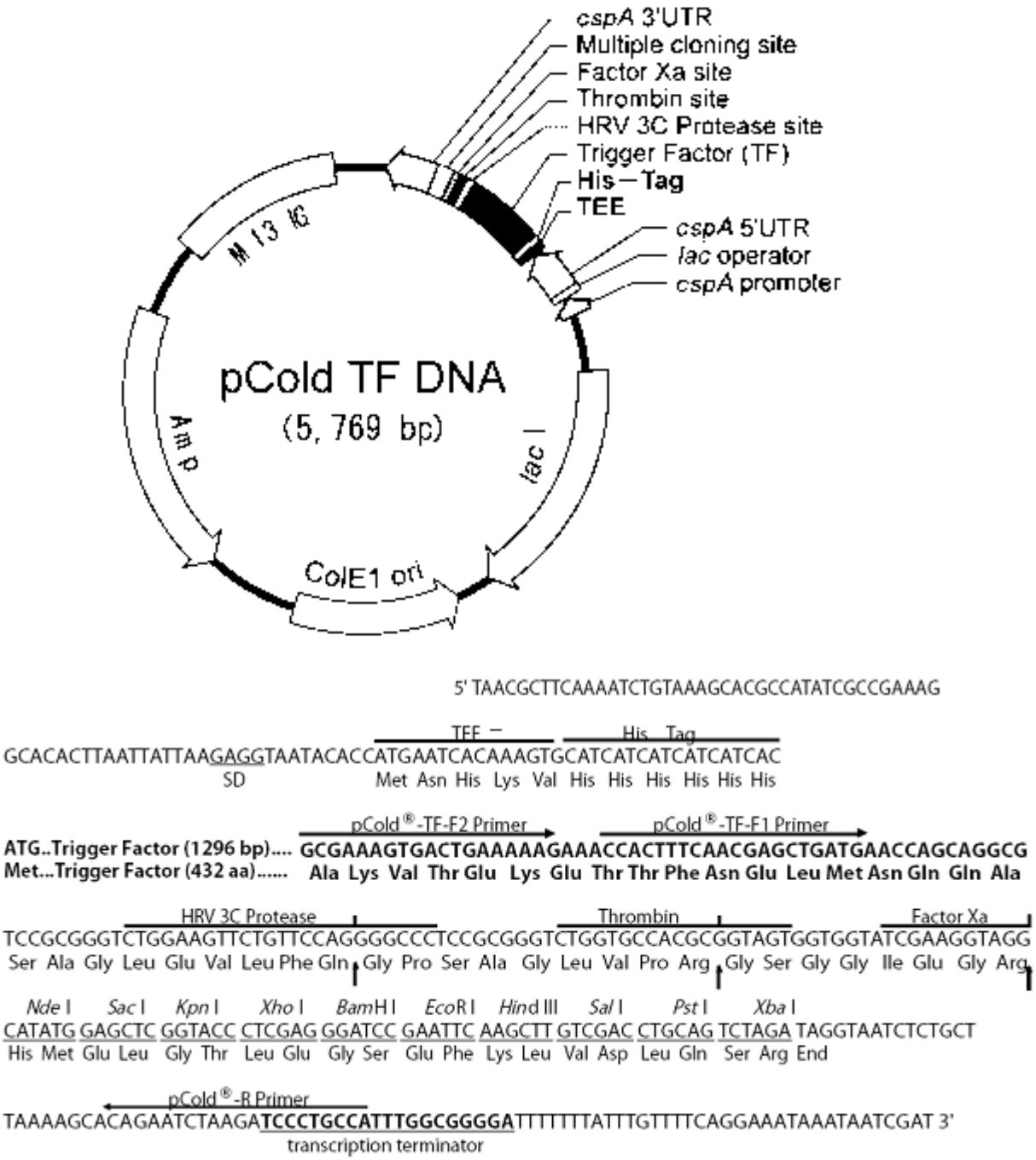


Figure 15. Vector map and MCS of pCold TF bacterial expression system (Takara Bio, USA).

The codon optimized CCHFV NP DNA gene was rescued from pUC57 cloning vector using the restriction sites (BamH1 and Pst1). Following restriction enzyme digestion as previously described (MATERIALS AND METHODS, 2.2.12) the

optimized CCHFV NP gene was subcloned into pCold TF generating pColdTF-opCCHFV NP. A ligation reaction to insert the opCCHFV NP gene into pCold TF bacterial expression vector was performed as described in Table 11. The ligation reaction was incubated for 1 hour at 16ºC.

Table 11. Ligation reaction for cloning CCHFV NP into pColdTF bacterial expression vector (Fermentas, USA).

Reaction component	Reaction volume
1 x T4 ligase buffer (400mM Tris-HCL, 100mM $MgCl_2$, 100mM DTT, 5mM ATP, pH 7.8)	1µl
T4 ligase enzyme (1U/µl)	1µl
Purified double digested CCHFV NP DNA (98.2ng)	6 µl
Purified double digested pColdTF plasmid (105.6ng)	2µl
Total	10µl

A 100µl aliquot of *E.coli* OverExpress C43 (DE3) competent cells prepared previously (MATERIALS AND METHODS, 2.2.8) were transformed with a 10µl aliquot of the ligation reaction. The transformation efficiency of the cells was 1 x 10^7 cfu/µg DNA. Colonies were selected from respective agar plates to identify positive transformants.

3.2.8 Plasmid purification

The recombinant plasmid was propagated as described previously (MATERIALS AND METHODS 2.2.11) and purified using the Pure Yield™ Plasmid Miniprep System (Promega, USA). All centrifugation steps were performed at 16000 x g. Briefly, 600µl of bacterial culture grown O/N in LB/amp media was transferred to a 1.5ml eppendorf. A 100µl aliquot of cell lysis buffer (supplied in the kit) and mixed by inverting the tube 6 times. A 350µl aliquot of cold (4-8°C) neutralization solution (supplied in the kit) was added and mixed by inverting the tube. The solution was centrifuged for 3 min. The supernatant was transferred to a PureYield™ minicolumn and the minicolum placed in a PureYield™ collection tube and centrifuged for 1 min. The flow through was discarded and placed in the same collection tube. A 200µl aliquot of endotoxin removal wash (supplied in the kit) was added to the

column and the column centrifuged for 1 min. A 400µl aliquot of column wash solution (supplied in the kit) was added to the minicolumn and the column centrifuged for 1min. The minicolumn was transferred to a clean 1.5ml eppendorf and 30µl of nuclease free water was added. The column was left to stand for 1 min and then centrifuged to elute the DNA. The purified DNA was stored at -20ºC. The recombinant plasmid was analyzed using gel electrophoresis and restriction enzyme site digestion (MATERIALS AND METHODS, 2.2.11) and the DNA concentration measured (MATERIALS AND METHODS, 2.2.6). A glycerol stock of the positive transformant was prepared (MATERIALS AND METHODS, 2.2.11). The control plasmid was prepared similarly.

3.2.9 DNA sequencing of pColdTF-opCCHFV NP

The pColdTF-opCCHFV NP DNA was sequenced as previously described (MATERIAL AND METHODS, 2.2.13). A pair of primers that target regions of the pColdTF plasmid and a pair of internal primers designed to obtain the complete codon optimized CCHFV NP gene were used. The primer sequences are shown in Table 12. The reaction mixtures were set up according to the recommendations described in Table 5 (MATERIALS AND METHODS, 2.2.13).

Table 12. Nucleotide sequences of primers used for sequencing DNA insert in pCold TF bacterial expression vector (Takara Bio, USA) and internal primers designed for sequencing CCHFV NP of SPU415/85 strain (J. Smith, unpublished data).

Sequencing primer name	Primer sequence	Genomic position	Target
pCold F1 (Forward)	5’ CCA CTT TCA ACG AGC TGA TG 3’	1153-1172	pCold TF
pCold R (Reverse)	5’ GGC AGG GAT CTT AGA TTC TG 3’	3482- 3463	pColdTF
SF2 (Forward internal)	5’ GGT TTC CGT GTC AAT GCA AAC 3’	409-429	CCHFV NP
SR3 (Reverse internal)	5’ CAT TGG GGT GCT CAG CAG AG 3’	1056-1037	CCHFV NP

3.2.10 Bacterial expression of pColdTF-opCCHFV NP using IPTG induction

An O/N culture of the recombinant pColdTF-opCCHFV NP was prepared from the glycerol stock (MATERIALS AND METHODS, 3.2.8). A 5ml volume of LB/amp was

inoculated with 10µl of a glycerol stock of pColdTF-opCCHFV NP. An additional 5ml volume of LB/amp was inoculated with 10µl glycerol stock of pCold TF with no gene insert as the negative control. The following day cultures were diluted 1:20 with pre-warmed LB/amp (2ml culture in 38ml LB/amp) and incubated for approximately 2 hours at 37ºC with shaking at 200rpm. The OD readings of the cells were measured with a spectrophotometer. When the OD_{600} reached between 0.4 - 0.5, the cultures were incubated at 16ºC for 30min. The cultures were then induced with IPTG to a final concentration of 1mM and incubated at 16ºC with shaking at 200rpm. Aliquots of 1ml of the induced cultures were collected at t = 0, t = 4 h and t = 24h post-induction. The cells were harvested from the samples by centrifugation at 16000 × g for 10min using the bench top centrifuge. The supernatant was discarded and the cells resuspended in PBS (MATERIALS AND METHODS, 3.2.3). The proteins were separated using an 8% resolving and 4% stacking SDS gel as described in MATERIALS AND METHODS, (3.2.4).

3.2.11 Protein solubility

A solubility study was performed to determine if the recombinant protein was present in the soluble or insoluble phase. Briefly, a 1ml sample of the pColdTF-opCCHFV NP culture was collected at t = 0, t = 4h and t = 24h. The samples were centrifuged at 16000 × g and the pellets were re-suspended in 500µl cold PBS. Sarcosyl (Sigma-Alrich, UK), a mild detergent that releases membrane bound protein was added to the samples at a final concentration of 7.5%. The cells were sonicated using a Branson 220 ultrasonic cleaner (SmithKline Company, USA) using 10 × 15 sec bursts with a 15 sec cooling period on ice between each burst. An 80µl aliquot of the cell suspension was collected to represent the total protein (soluble and insoluble fraction) and added to 20µl of 5 × protein loading buffer. Another 80µl aliquot of sample was collected from the suspension and centrifuged at 16000 × g for 20min. The supernatant was collected (soluble fraction) and added to 20µl of the loading buffer. The pellet (insoluble fraction) was re-suspended in 500µl cold PBS, 80µl was collected and added to 20µl loading buffer. All the samples were heated for 5min at 95ºC and separated by electrophoresis on an SDS PAGE gel (8% resolving, 4% stacking) as previously described (MATERIALS AND METHODS, 3.2.4).

3.2.12 Denaturation, purification and refolding of recombinant His-tagged proteins from the insoluble phase

The recombinant NP protein has a 5' 6×His tag which was used to purify the protein from other *E. coli* proteins under denaturing conditions using Protino®Ni-TED resin according to the manufacturer's instructions (Machery-Nagel, USA). Protino®Ni-TED products enable fast and convenient purification of recombinant polyhistidine-tagged proteins by immobilized metal ion affinity chromatography (IMAC). Protino®Ni-TED is a dry silica-based resin precharged with Ni^{2+} ions. Protein binding is based on the interaction between the polyhistidine tag of the recombinant protein and immobilized Ni^{2+} ions. The chelating group of Protino® Ni-TED is based on TED (tris-carboxymethyl ethylene diamine), a strong pentadentate metal chelator. TED occupies five of the six binding sites in the coordination sphere of the Ni^{2+} ion. Compared to TED, other chelating groups such as NTA (nitrilotriacetic acid) have four binding sites available for the Ni^{2+} ion. The additional chelation site of TED with Ni^{2+} minimizes metal leaching during purification and increases specificity for polyhistidine-tagged proteins. As a result target protein of excellent purity is eluted from the column.

For purification of larger quantities of protein to obtain sufficient for ELISA antigen, a scale up to 250ml culture was performed. An O/N culture of the recombinant pColdTF-opCCHFV NP was prepared from the glycerol stock (METHODS AND MATERIALS, 3.2.8). A 25ml volume of LB/amp was inoculated with 250µl of a glycerol stock of pColdTF-opCCHFV NP. The following day the culture was diluted 1:10 with prewarmed LB/amp (25ml culture in 225ml LB/amp) and incubated for approximately 2 hours at 37ºC with shaking at 200rpm. The cultures were induced with IPTG as previously described (MATERIALS AND METHODS, 3.2.10) and allowed to shake for 24h before harvesting cells by centrifugation at 6000 × g for 10 min at 4ºC. The cells were lysed by resuspending the pellet in Bugbuster Protein Extraction Reagent (Novagen, USA) at a final concentration of 200mg/ml. An aliquot of r-lysozyme (Novagen, USA) was added to achieve a final concentration of 1mg/ml and 50 units/ml of benzonase (Novagen, USA) was added to the cell suspension. The cell suspension was stirred at 22-25ºC for 30 min. The cells were then sonicated on ice using 10 × 15 sec bursts with a 15 sec cooling period on ice

between each burst. After sonication the cells were centrifuged at 10000 × g for 30 min at 4ºC. The pellet were re-suspended in 2.5ml 1× lysis equilibrium buffer (LEW) buffer (supplied in the kit). The re-suspended samples were centrifuged at 10000 × g for 30 min at 4ºC. The pellet was re-suspended in 2.5ml of denaturing solubilization (DS) buffer (pH 8.0) containing 8M urea, pooled and stirred on ice for 60 min. The cell suspension was centrifuged at 10000 × g for 30 min at 20ºC. The supernatant was transferred to a clean tube. Two Protino nickel columns were equilibrated with 2ml DS buffer and allowed to drain by gravity. The supernatant was added to the column and the column allowed to drain by gravity. The column was washed with 2 × 2ml DS buffer. The protein was eluted with 4.5ml of denaturing elution buffer containing 8M urea and 0.25M imidazole and the eluates collected in 500µl aliquots.

To refold the protein after denaturation the protein eluates were diluted 1:1 in LEW buffer (500µl protein + 500µl LEW buffer) and ultra-filtered to a final volume of approximately 500µl per eluate using Millipore filters with a 30 kDa molecular weight cut off (Millipore Corporation, USA). The control antigen (pColdTF with no gene insert) was prepared similarly. The purified proteins were separated on SDS PAGE (8% resolving and 4% stacking) to confirm the purity of the proteins and were stored at -70ºC.

3.2.13 Concentration of proteins

The protein concentrations were measured using the Quant-iT Protein Assay Kit (Invitrogen, USA) according to the manufacturer's instructions. The Quant-iT Protein Assay is supplied with three standards. A Quant-iT working solution was prepared by diluting the Quant-iT Protein reagent (supplied in the kit) 1:200 in Quant-iT Protein buffer (supplied in the kit). An aliquot of 190µl of Quant-iT working solution was loaded into each of the tubes used for the standards. A 10µl aliquot of each Quant-iT standard (supplied in the kit) was added to the appropriate tube and mixed by vortexing for 2-3 seconds. Protein samples diluted 1:200 in Quant-iT working solution were briefly mixed by vortexing for 2-3 sec. All tubes were allowed to incubate at room temperature for 15 min. The Qubit fluorometer was calibrated for each measure using the three standard solutions prepared. The

reading was recorded in µg/ml and the dilution was considered when calculating the final concentration. The proteins were separated using an 8% resolving and 4% stacking SDS gel respectively as described in MATERIALS AND METHODS, (3.2.4).

3.2.14 Characterization of expressed His tagged opCCHFV NP by Western blot analysis

The Western blot analysis was performed using the Pico Fast Western Blot Kit (Pierce Biotechnology, USA). The proteins were transferred from the gel to a PVDF membrane as described in MATERIALS AND METHODS, (3.2.5). The membrane was briefly washed in fast western 1x wash buffer (supplied in the kit) to remove transfer buffer. A mouse anti-His IgG antibody (100µg/ml) (Roche, Germany) diluted 1:200 in the fast western antibody diluent (supplied in the kit) was added to the membrane. The membrane was incubated for 30min at 22-25ºC with shaking. The membrane was removed from the anti-HIS antibody solution and placed in a clean incubation tray. Anti-mouse IgG horse radish peroxidase (HRP) reagent (supplied in the kit) diluted 1:500 in antibody diluents was added to the membrane and the membrane incubated for 10min with shaking at 22-25ºC. The membrane was removed from the HRP solution and washed by suspension in approximately 20ml of 1x wash buffer, agitating it for 5 min. The wash was repeated 3 times. The membrane was then exposed to the supersignal west pico solution (supplied in the kit) mixed according to the manufacturers instructions for 1 min at 22-25ºC. The membrane was placed in a clear plastic wrap and exposed to film (Thermo Scientific, USA) for approximately 60 sec. The film was incubated for 2 min in developer and for 1 min in fixer. The film was washed with water to remove excess fixer and visualized.

3.3 Results

3.3.1 Cloning of the CCHFV nucleoprotein gene in pQE-80L bacterial expression vector

The confirmation of insertion of the correct DNA into pQE-80L vector is shown in Figure 16. The undigested supercoiled pQE-80L plasmid DNA (lane 2) migrated faster than the linearized recombinant pQE-80L-CCHFV NP DNA (lane 3). The pQE-80L-CCHFV NP DNA which was linearized with BamH1 and Pst1 yielded the expected 1443 bp fragment. The DNA concentration of the digested pQE-80L-CCHFV NP was 68.7ng/µl.

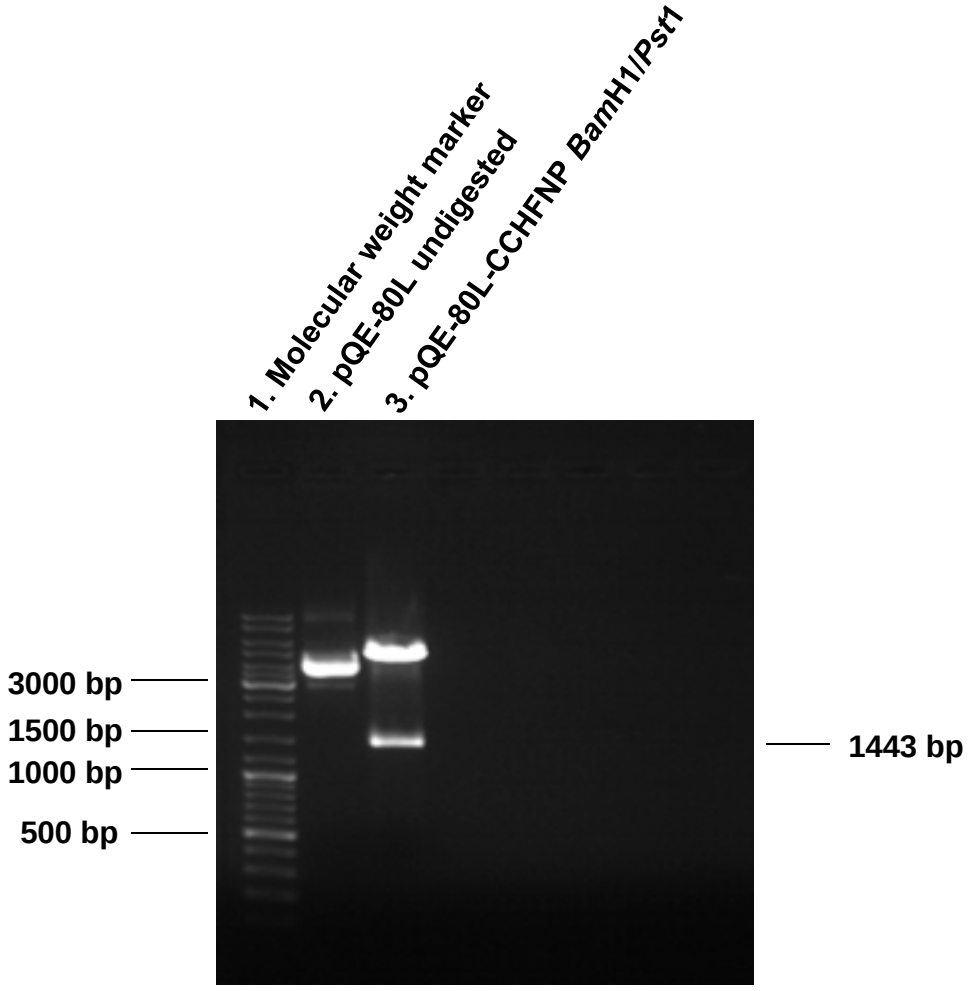


Figure 16. Agarose gel electrophoretic analysis of restriction enzyme analysis of a plasmid obtained from the ligation of the CCHFV NP gene into pQE-80L. Lane 1- O'GeneRuler DNA ladder mix molecular weight marker (Fermentas, USA); Lane 2 – Undigested pQE-80L; Lane 3 – pCCHFV NP-80L digested with BamH1 and Pst1.

3.3.2 Sequencing and sequence analysis of the gene encoding CCHFV SPU415/85 in pQE-80L vector

The nucleotide sequence of CCHFV NP gene 3 determined in pQE-80L is shown in Figure 17. Analysis of the sequence data verified insertional cloning of the correct DNA fragment of 1443bp. The S segment ORF is in frame with the start codon, the RGS his epitope, the His tag sequence and the stop codon on the pQE-80L plasmid as indicated in Figure 17. The nucleotide blast performed on the edited sequence confirmed the identity of the gene as CCHFV.

Figure 17. Nucleotide sequence of CCHFV NP gene 3 sequenced in pQE-80L. The ATG start codon and TGA stop codon of pQE-80L are highlighted in **grey**. The RGS His epitope is highlighted in **pink**. The His_6 tag sequence of pQE-80L is highlited in **turquoise**. The forward and reverse RT-PCR primers of CCHFV NP are highlighted in **yellow** with BamH1 and Pst1 restriction sites higlighted in **red** on the forward and reverse primer respectively.

CCHFV NP3 SPU415/85 (pQE-80L)

```
1                                                                    60
ATGAGAGGATCGCATCACCATCACCATCACGGATCCGAAAACAAAATTGAGGTGAATAAC

61                                                                   120
AAAGATGAGATGAACAAGTGGTTTGAAGAGTTCAAAAAAGGAAATGGACTTGTGGACACC

121                                                                  180
TTCACAAACTCCTATTCCTTTTGTGAGAGTGTTCCAAATTTGGACAAGTTTGTGTTCCAA

181                                                                  240
ATGGCCAGTGCCACCGATGATGCACAAAAGGATTCCATCTACGCGTCTGCTCTGGTGGAG

241                                                                  300
GCAACAAAATTTTGTGCACCTGTATATGAGTGTGCATGGGTTAGCTCCACTGGCATTGTG

301                                                                  360
AAAAGGGGACTTGAATGGTTCGAAAAAAATGCGGGCACCATTAAGTCTTGGGATGAAAGT

361                                                                  420
TATACTGAGCTAAAAGTCGACGTCCCAAAAATAGAACAGCTAGCCAATTACCAACAGGCT

421                                                                  480
GCCTTGAAATGGAGGAAAGACATAGGTTTCCGTGTCAATGCAAACACAGCGGCTCTGAGT

481                                                                  540
AACAAAGTCCTCGCAGAGTACAAAGTTCCTGGAGAGATTGTGATGTCTGTCAAAGAGATG

541                                                                  600
CTGTCAGACATGATTAGGAGAAGGAACCAGATTCTAAACAGGGGTGGTGATGAGAATCCA

601                                                                  660
CGTGGCCCTGTGAGCCGTGAGCATGTAGACTGGTGCAGGGAGTTTGTCAAAGGCAAATAC

661                                                                  720
ATCATGGCCTTCAACCCACCATGGGGGGACATCAACAAGTCAGGCCGTTCAGGAATAGCA
```

721 780
CTTGTCGCAACAGGCCTTGCCAAGCTTGCAGAGACTGAGGGAAAGGGAGTATTTGACGAA

781 840
GCCAAAAAAGACTGTGGAAGCCCTCAATGGGTATCTGGACAAGCACAAGGACGAAGTTGAC

841 900
AGAGCGAGTGCTGACAGCATGATAACAAACCTTCTCAAGCACATTGCTAAGGCACAGGAG

901 960
CTTTATAAGAATTCGTCTGCACTCCGTGCACAAGGTGCACAGATTGACACTGCTTTCAGC

961 1020
TCATACTATTGGCTTTACAAGGCTGGCGTGACCCCAGAGACATTCCCGACGGTGTCACAG

1021 1080
TTCCTCTTTGAGCTAGGAAAACAGCCAAGAGGTACCAAGAAAATGAAGAAGGCTCTGCTG

1081 1140
AGCACCCCAATGAAGTGGGGGAAGAAACTTTATGAACTCTTTGCCGACGATACTTTCCAG

1141 1200
CAGAACAGGATCTACATGCACCCTGCCGTGCTTACAGCTGGCAGAATCAGTGAAATGGGA

1201 1260
GTCTGCTTTGGGACAATCCCCGTGGCCAACCCTGATGATGCTGCTCAAGGGTCTGGACAT

1261 1320
ACCAAGTCTATTCTCAACCTCCGGACTAACACCGAGACCAACAATCCGTGTGCCAGGACC

1321 1380
ATTGTCAAGCTGTTTGAAATTCAGAAAACAGGGTTCAACATTCAGGACATGGACATAGTG

1381 1440
GCCTCTGAGCACTTGCTG**CAT**CAGTCTCTTGTTGGCAAGCAATCTCCATTCCAGAACGCC

1441 1488
TACAACGTCAAGGG**CAATGCCACCAGTGTTAACATTATCCTGCAGTGA**

3.3.3 Bacterial expression of pQE-80L-CCHFV NP using IPTG induction

Protein fractions collected at t=0, t = 1h, t = 3 and t = O/ N were separated on an 8% resolving gel and 4% stacking gel as shown in Figure 18. The expected size of the recombinant CCHFV NP was 54kDa. As shown on Figure 18 there was no detectable protein of 54 kDA from the CCHFV construct. Attempts to express the CCHFV NP protein by optimizing IPTG concentrations and temperature conditions did not result in expression of a 54kDa protein. Western blot analysis of the protein gel shown in Figure 19 confirmed that no protein with a His tag could be detected.

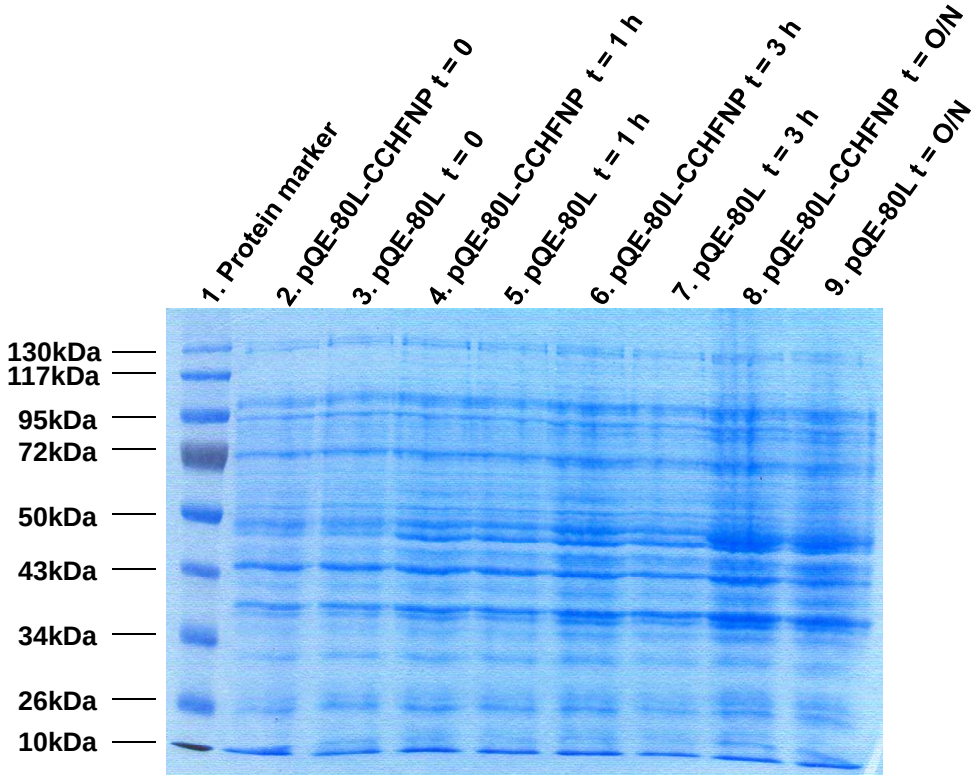


Figure 18. SDS PAGE analysis of proteins expressed using the pQE-expression vector, CCHFV NP in JM109 competent cells.

Lanes each contain 25µl of sample: Lane 1 – (Prestained Protein Ladder comprising proteins from 10 to 130kDa (Fermentas, USA); Lane 2 – pQE-80L-CCHFV NP t = 0; Lane 3 – pQE-80L negative control expression t = 0; Lane 4 – pQE-80L-CCHFV NP t = 1 hr; Lane 5 – pQE-80L negative control expression t = 1 h; Lane 6 – pQE-80L-CCHFV NP t = 3 h; Lane7 – pQE-80L negative control expression t = 3 h; Lane 8 - pQE-80L-CCHFV NP t = O/N; Lane 9 - pQE-80L negative control expression t = O/N.

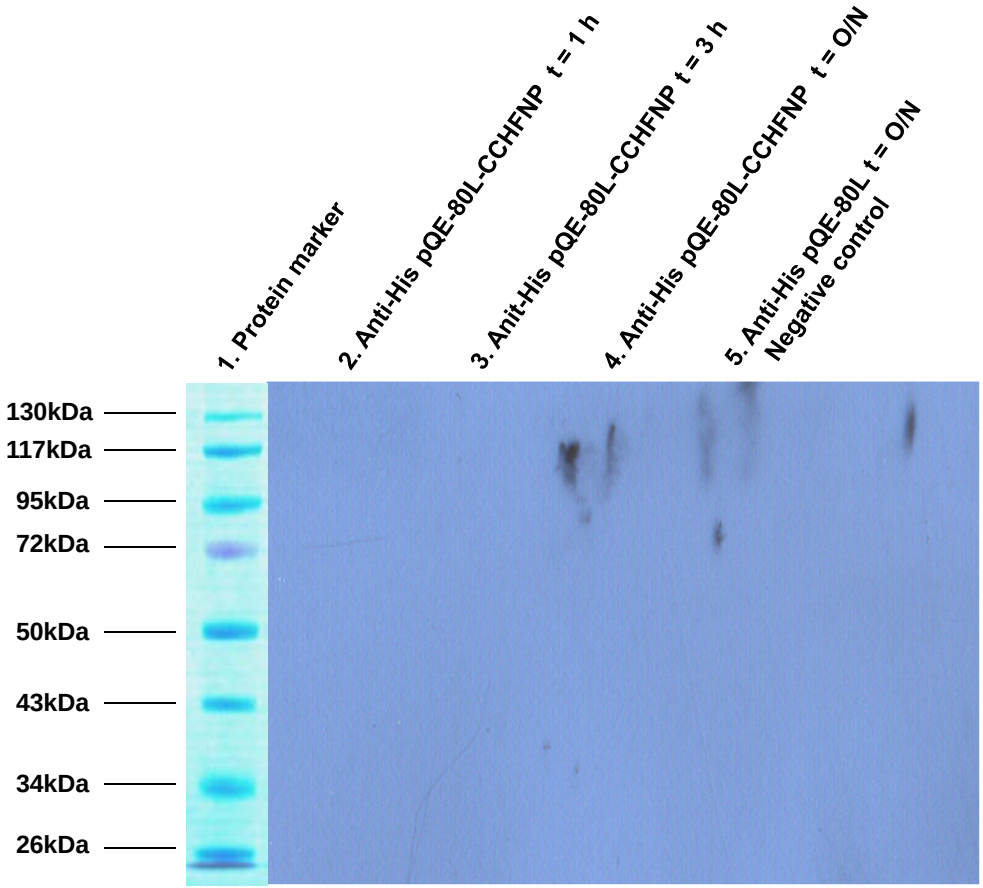


Figure 19. Western blot analysis of attempted expression of CCHFV NP protein in pQE-80L

L: Lane 1 – (Prestained Protein Ladder comprising proteins from 10 to 130kDa (Fermentas, USA); Lane 2 – Anti-His pQE-80L-CCHFV NP t = 1h; Lane 3 – pQE-80L-CCHFV NP t = 3h; Lane 4 - pQE-80L-CCHFV NP t = O/N; Lane 5 – Anti-His pQE-80L negative control t = O/N.

3.3.4 Codon optimization of CCHFV NP gene

A wide variety of factors regulate and influence gene expression levels, the OptimumGene algorithm takes into consideration GC content, secondary mRNA structures, poly A sites and RNA instability, producing the single gene that can reached the highest possible level of expression. In this case, our CCHFV native gene employed tandem rare codons that reduced the efficiency of translation or disengaged the translational machinery.

The two indices that measure codon bias are CAI and Enc are used to measure the optimization process. CAI is the most effective of all codon bias measures for predicting gene expression levels and measures the similarity between the codon usage of a gene and the codon usage of the host.

Figure 20 demonstrates the distribution of codon frequency usage along the length of the gene sequence of the S segment of CCHFV SPU 415/85 before optimization. The frequency at which each codon is utilized by the *E.coli* host is expressed as a percentage with 100% being optimal usage. The codon usage for the NP gene of CCHFV ranged from 10% to 70% along the length of the gene. The CAI was calculated as 0.64 indicating that several of the codons used by the native CCHFV gene were less frequently recognized by *E.coli.*

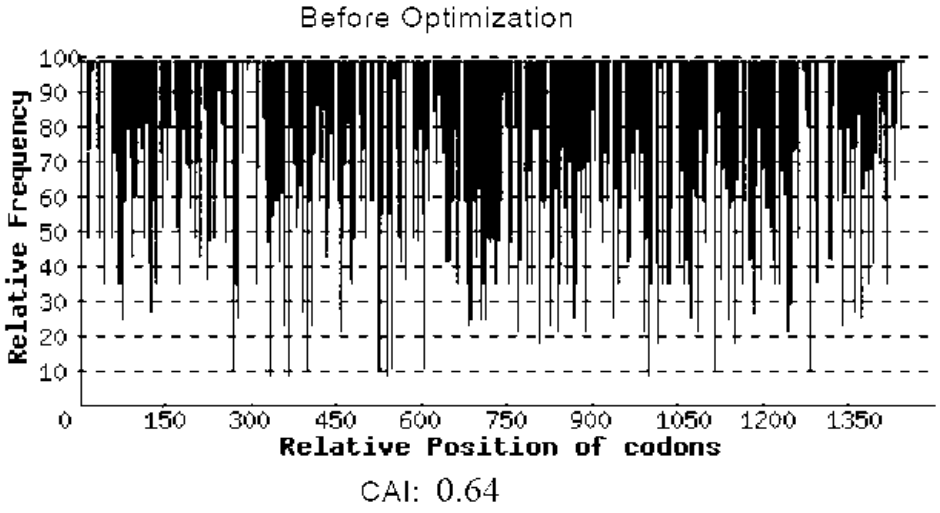


Figure 20. The distribution of codon usage frequency along the length of the gene sequence of the S segment of CCHFV SPU 415/85 before optimization.

After codon optimization the codon usage of *E. coli* along the length of the CCHFV NP gene significantly increased with values ranging from approximately 70% to 100%, with the exception of four codons with a 40-50% usage. The CAI of the codon optimized gene was calculated as 0.95 as demonstrated in Figure 21.

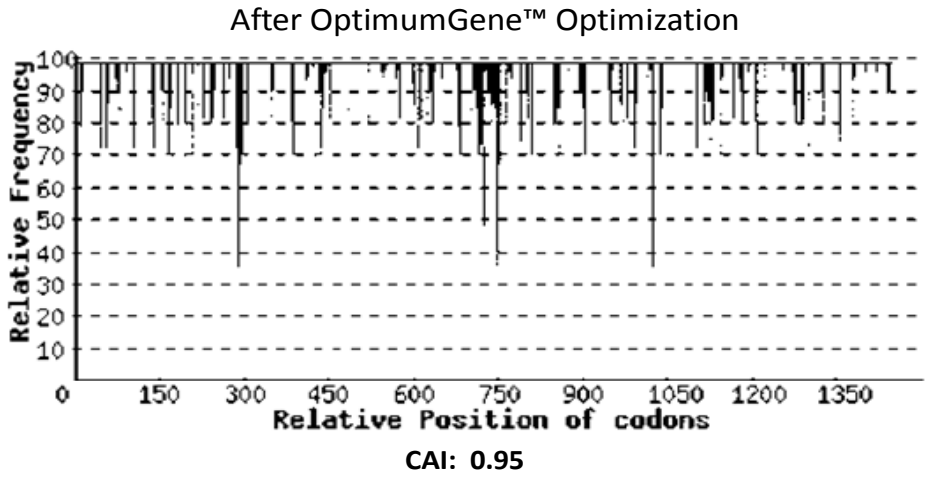


Figure 21. The distribution of codon usage frequency along the length of the gene of the S segment of CCHFV SPU 415/85 sequence after optimization.

To further illustrate the codon usage Figure 22 shows the percentage distribution from the CCHFV NP gene recognized by *E.coli*. The codons have been grouped according to usage by *E.coli* where 0-10% is the group less frequently recognized and 91-100% is the group most frequently utilized. Less than half of the codons (42%) are frequently utilized by *E.coli.* The proportion of codons utilized by *E.coli* was increased to 83% after codon optimization (Figure 23).

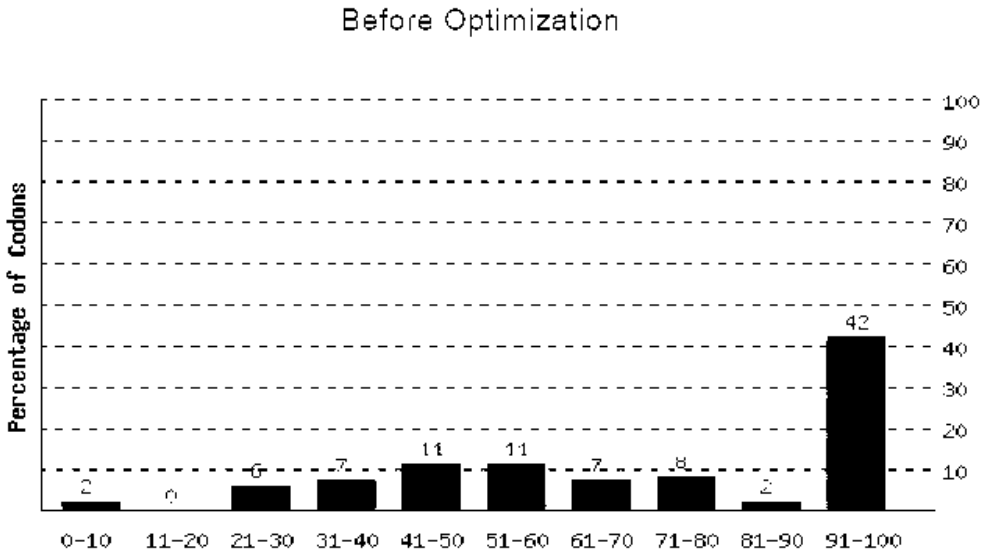


Figure 22. The percentage distribution of codons in computed codon quality groups before optimization.

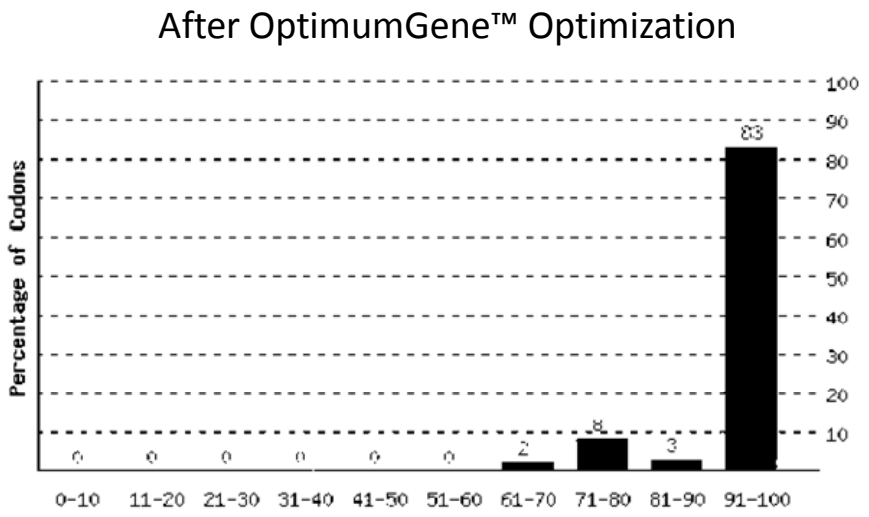


Figure 23. The percentage distribution of codons in computed codon quality groups after optimization.

GC content and unfavorable peaks were optimized to stabilize the mRNA. The average GC content as shown in Figure 24 was 47.78% before optimization and increased to 49.89 after optimization (Figure 25). The Stem-Loop structures, which impact ribosomal binding and stability of mRNA, were broken.

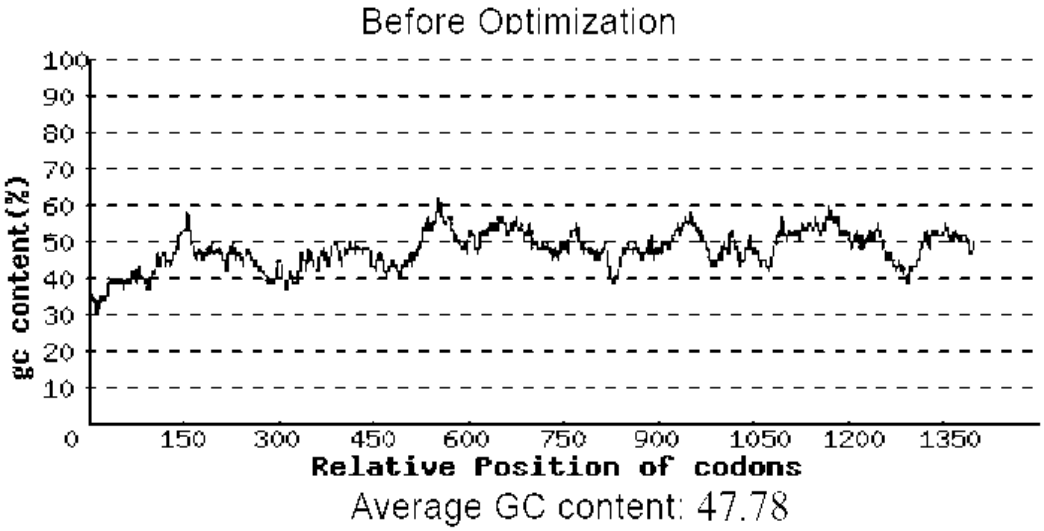
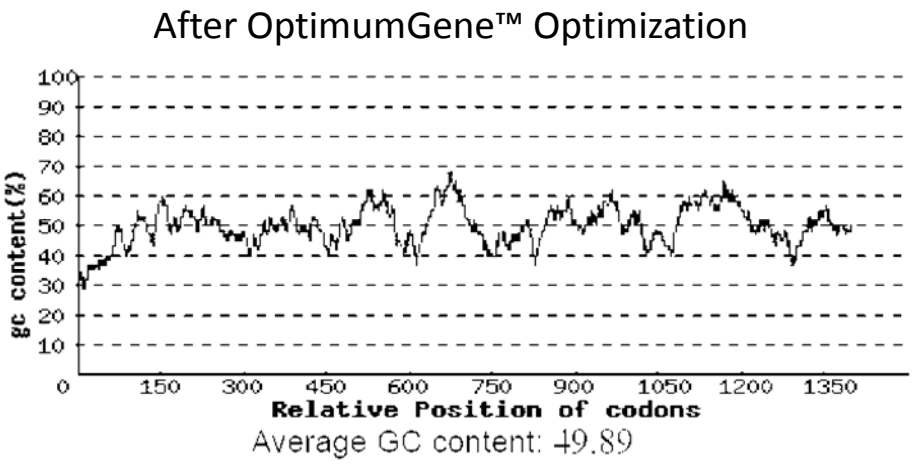


Figure 24. Average GC content before optimization.



Figue 25. Average GC content after optimization.

The nucleotide sequence of the native gene encoding the NP of CCHFV and the codon optimized gene were aligned and translated to show that the predicted amino acid sequence was identical for both coding regions (Figure 26).

Figure 26. The nucleotide sequence and predicted amino acid sequence of native (original) and codon-optimized CCHFV NP gene SPU 415/85 (amino acid residues 1 to 481). The optimized codons are shaded in **grey**.

```
          1                                                          60
optimized GAAAACAAAATCGAAGTGAACAACAAAGATGAAATGAACAAATGGTTCGAAGAATTCAAA
original  GAAAACAAAATTGAGGTGAATAACAAAGATGAGATGAACAAGTGGTTTGAAGAGTTCAAA
          E  N  K  I  E  V  N  N  K  D  E  M  N  K  W  F  E  E  F  K

          61                                                        120
optimized AAAGGCAACGGTCTGGTTGATACCTTTACCAATAGCTATAGCTTCTGCGAAAGCGTGCCG
original  AAAGGAAATGGACTTGTGGACACCTTCACAAACTCCTATTCCTTTTGTGAGAGTGTTCCA
          K  G  N  G  L  V  D  T  F  T  N  S  Y  S  F  C  E  S  V  P

          121                                                       180
optimized AACCTGGATAAATTTGTTTTCCAGATGGCGAGCGCCACCGATGATGCACAGAAAGATAGC
original  AATTTGGACAAGTTTGTGTTCCAAATGGCCAGTGCCACCGATGATGCACAAAAGGATTCC
          N  L  D  K  F  V  F  Q  M  A  S  A  T  D  D  A  Q  K  D  S

          181                                                       240
optimized ATCTATGCGAGCGCCCTGGTGGAAGCAACCAAATTTTGTGCGCCGGTTTATGAATGCGCC
original  ATCTACGCGTCTGCTCTGGTGGAGGCAACAAAATTTTGTGCACCTGTATATGAGTGTGCA
          I  Y  A  S  A  L  V  E  A  T  K  F  C  A  P  V  Y  E  C  A

          241                                                       300
optimized TGGGTTAGCAGCACCGGCATTGTTAAACGTGGTCTGGAATGGTTCGAAAAGAACGCAGGC
original  TGGGTTAGCTCCACTGGCATTGTGAAAAGGGGACTTGAATGGTTCGAAAAAAATGCGGGC
          W  V  S  S  T  G  I  V  K  R  G  L  E  W  F  E  K  N  A  G

          301                                                       360
optimized ACCATCAAAAGCTGGGATGAAAGCTATACCGAACTGAAAGTGGATGTTCCGAAAATTGAA
original  ACCATTAAGTCTTGGGATGAAAGTTATACTGAGCTAAAAGTCGACGTCCCAAAAATAGAA
          T  I  K  S  W  D  E  S  Y  T  E  L  K  V  D  V  P  K  I  E
```

```
          361                                                       420
optimized CAGCTGGCGAACTATCAGCAGGCCGCACTGAAATGGCGCAAAGATATCGGTTTTCGTGTG
original  CAGCTAGCCAATTACCAACAGGCTGCCTTGAAATGGAGGAAAGACATAGGTTTCCGTGTC
          Q  L  A  N  Y  Q  Q  A  A  L  K  W  R  K  D  I  G  F  R  V

          421                                                       480
optimized AATGCGAACACCGCCGCACTGAGCAATAAAGTTCTGGCGGAATATAAAGTGCCGGGCGAA
original  AATGCAAACACAGCGGCTCTGAGTAACAAAGTCCTCGCAGAGTACAAAGTTCCTGGAGAG
          N  A  N  T  A  A  L  S  N  K  V  L  A  E  Y  K  V  P  G  E

          481                                                       540
optimized ATTGTTATGAGCGTGAAAGAAATGCTGAGCGATATGATCCGCCGTCGCAACCAGATTCTG
original  ATTGTGATGTCTGTCAAAGAGATGCTGTCAGACATGATTAGGAGAAGGAACCAGATTCTA
          I  V  M  S  V  K  E  M  L  S  D  M  I  R  R  R  N  Q  I  L

          541                                                       600
optimized AATCGTGGTGGCGATGAAAACCCGCGCGGTCCGGTTAGCCGTGAACATGTGGATTGGTGT
original  AACAGGGGTGGTGATGAGAATCCACGTGGCCCTGTGAGCCGTGAGCATGTAGACTGGTGC
          N  R  G  G  D  E  N  P  R  G  P  V  S  R  E  H  V  D  W  C

          601                                                       660
optimized CGCGAATTCGTTAAAGGCAAATATATCATGGCCTTTAATCCGCCGTGGGGTGATATTAAC
original  AGGGAGTTTGTCAAAGGCAAATACATCATGGCCTTCAACCCACCATGGGGGGACATCAAC
          R  E  F  V  K  G  K  Y  I  M  A  F  N  P  P  W  G  D  I  N

          661                                                       720
optimized AAAAGCGGCCGTAGCGGTATCGCACTGGTGGCGACCGGCCTGGCCAAACTGGCAGAAACC
original  AAGTCAGGCCGTTCAGGAATAGCACTTGTCGCAACAGGCCTTGCCAAGCTTGCAGAGACT
          K  S  G  R  S  G  I  A  L  V  A  T  G  L  A  K  L  A  E  T

          721                                                       780
optimized GAGGGTAAAGGCGTTTTCGATGAAGCAAAGAAAACCGTGGAAGCCCTGAATGGTTATCTG
original  GAGGGAAAGGGAGTATTTGACGAAGCCAAAAAGACTGTGGAAGCCCTCAATGGGTATCTG
          E  G  K  G  V  F  D  E  A  K  K  T  V  E  A  L  N  G  Y  L
```

```
          781                                                        840
optimized GATAAACACAAAGATGAAGTTGATCGCGCAAGCGCGGATAGCATGATTACCAACCTGCTG
original  GACAAGCACAAGGACGAAGTTGACAGAGCGAGTGCTGACAGCATGATAACAAACCTTCTC
          D  K  H  K  D  E  V  D  R  A  S  A  D  S  M  I  T  N  L  L

          841                                                        900
optimized AAACATATCGCCAAAGCACAGGAACTGTATAAAAATAGCAGCGCGCTGCGTGCCCAGGGC
original  AAGCACATTGCTAAGGCACAGGAGCTTTATAAGAATTCGTCTGCACTCCGTGCACAAGGT
          K  H  I  A  K  A  Q  E  L  Y  K  N  S  S  A  L  R  A  Q  G

          901                                                        960
optimized GCACAGATTGATACCGCGTTTAGCAGCTATTATTGGCTGTATAAAGCCGGTGTGACCCCG
original  GCACAGATTGACACTGCTTTCAGCTCATACTATTGGCTTTACAAGGCTGGCGTGACCCCA
          A  Q  I  D  T  A  F  S  S  Y  Y  W  L  Y  K  A  G  V  T  P

          961                                                       1020
optimized GAAACCTTCCCGACCGTTAGCCAGTTTCTGTTCGAACTGGGCAAACAGCCGCGCGGTACC
original  GAGACATTCCCGACGGTGTCACAGTTCCTCTTTGAGCTAGGAAAACAGCCAAGAGGTACC
          E  T  F  P  T  V  S  Q  F  L  F  E  L  G  K  Q  P  R  G  T

          1021                                                      1080
optimized AAGAAAATGAAAAAAGCACTGCTGAGCACCCCGATGAAATGGGGCAAAAAACTGTATGAA
original  AAGAAAATGAAGAAGGCTCTGCTGAGCACCCCAATGAAGTGGGGGAAGAAACTTTATGAA
          K  K  M  K  K  A  L  L  S  T  P  M  K  W  G  K  K  L  Y  E

          1081                                                      1140
optimized CTGTTTGCGGATGATACCTTCCAGCAGAACCGTATCTATATGCACCCGGCCGTGCTGACC
original  CTCTTTGCCGACGATACTTTCCAGCAGAACAGGATCTACATGCACCCTGCCGTGCTTACA
          L  F  A  D  D  T  F  Q  Q  N  R  I  Y  M  H  P  A  V  L  T

          1141                                                      1200
optimized GCAGGTCGCATTAGCGAAATGGGCGTTTGCTTTGGTACCATCCCGGTGGCGAATCCGGAT
original  GCTGGCAGAATCAGTGAAATGGGAGTCTGCTTTGGGACAATCCCCGTGGCCAACCCTGAT
          A  G  R  I  S  E  M  G  V  C  F  G  T  I  P  V  A  N  P  D
```

```
          1201                                                        1260
optimized GATGCCGCACAGGGCAGCGGTCATACCAAAAGCATTCTGAACCTGCGTACCAATACCGAA
original  GATGCTGCTCAAGGGTCTGGACATACCAAGTCTATTCTCAACCTCCGGACTAACACCGAG
          D  A  A  Q  G  S  G  H  T  K  S  I  L  N  L  R  T  N  T  E

          1261                                                        1320
optimized ACCAACAATCCGTGTGCCCGCACCATCGTTAAACTGTTCGAAATTCAGAAAACCGGCTTT
original  ACCAACAATCCGTGTGCCAGGACCATTGTCAAGCTGTTTGAAATTCAGAAAACAGGGTTC
          T  N  N  P  C  A  R  T  I  V  K  L  F  E  I  Q  K  T  G  F

          1321                                                        1380
optimized AACATCCAGGATATGGATATTGTGGCGAGCGAACACCTGCTGCATCAGAGCCTGGTTGGT
original  AACATTCAGGACATGGACATAGTGGCCTCTGAGCACTTGCTGCACCAGTCTCTTGTTGGC
          N  I  Q  D  M  D  I  V  A  S  E  H  L  L  H  Q  S  L  V  G

          1381                                                        1440
optimized AAACAGAGCCCGTTTCAGAATGCGTATAACGTGAAAGGCAATGCGACCAGCGTGAACATC
original  AAGCAATCTCCATTCCAGAACGCCTACAACGTCAAGGGCAATGCCACCAGTGTTAACATT
          K  Q  S  P  F  Q  N  A  Y  N  V  K  G  N  A  T  S  V  N  I

          1443
optimized ATT
original  ATC
          I
```

3.3.5 Cloning of gene encoding the codon optimized NP from pUC57 into pCold TF bacterial expression vector

Figure 27 demonstrates that restriction enzyme analysis of the pUC57-opCCHFV NP plasmid yielded the expected 1443 bp fragment (lane 3) after linearization with BamH1 and Pst1 confirming the presence of the opCCHFV NP gene in pUC57 vector. The DNA concentration of purified, gel extracted opCCHFV NP was 125.8ng/µl.

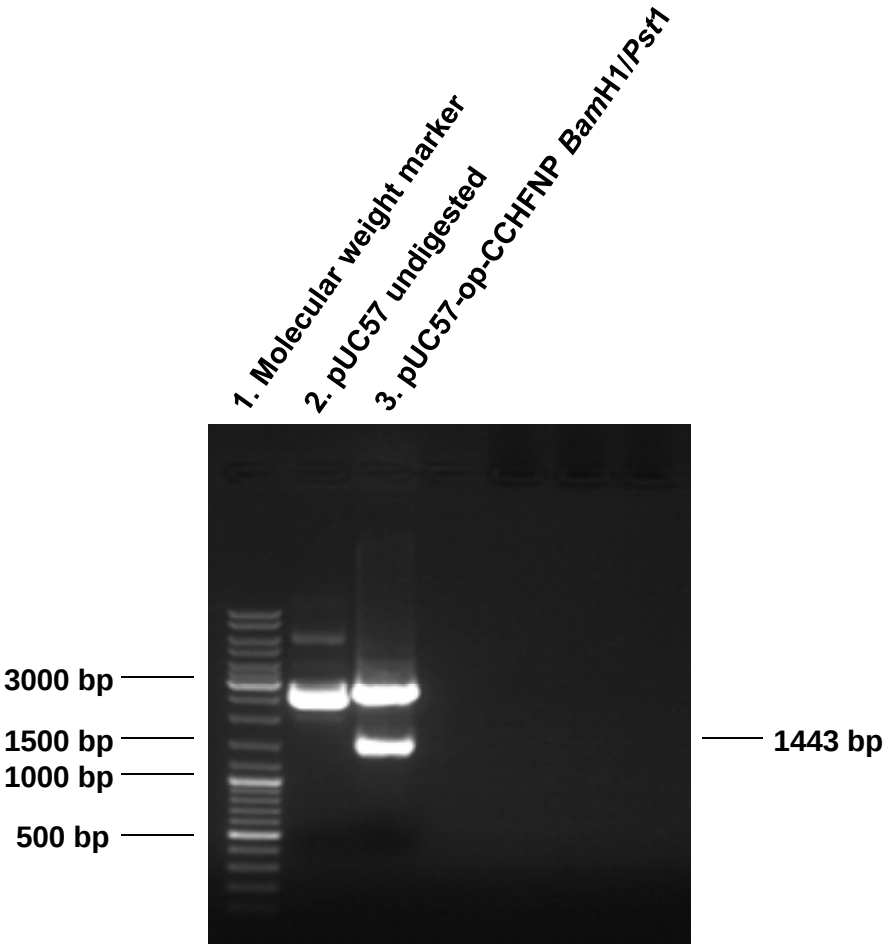


Figure 27. Agarose gel electrophoretic analysis of restriction enzyme analysis of a plasmid obtained from pUC57 vector supplied by GenScript. Lane 1 – O'GeneRuler DNA ladder mix molecular weight marker molecular weight marker; Lane 2 - 5µl of undigested pUC57; Lane 3 - 20µl of pUC57-opCCHFV NP digested with BamH1 and Pst1.

A 5µl aliquot of the purification was analyzed by gel electrophoresis on a 1% gel as shown in Figure 28. A single band of predicted size was present in lane 2.

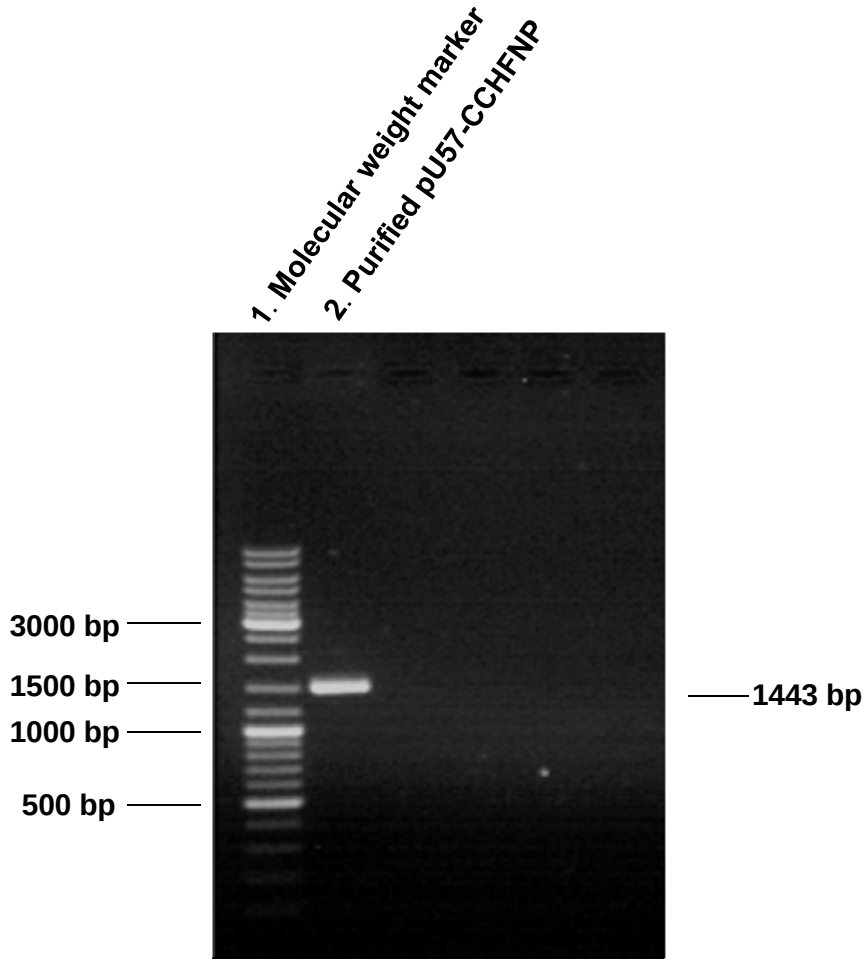


Figure 28. Agarose gel electrophoretic analysis of NP gene plasmid purification with QIAprep Spin Miniprep kit (QIAgen USA). Lane 1 – O'GeneRuler DNA ladder mix molecular weight marker molecular weight marker; Lane 2 - 5µl of 30µl clean-up reaction

A ligation reaction was performed to ligate opCCHFV NP into pColdTF vector using T4 DNA ligase as described in MATERIALS AND METHODS, (2.2.9) generating pColdTF-opCCHFV NP. Transformation of *E.coli* OverExpress C43 (DE3) competent cells with the recombinant pColdTF-opCCHFV NP plasmid DNA, purification and restriction enzyme analysis was performed as described (METHODS AND MATERIALS, 2.2.10, 3.2.7 and 2.2.12 respectively).

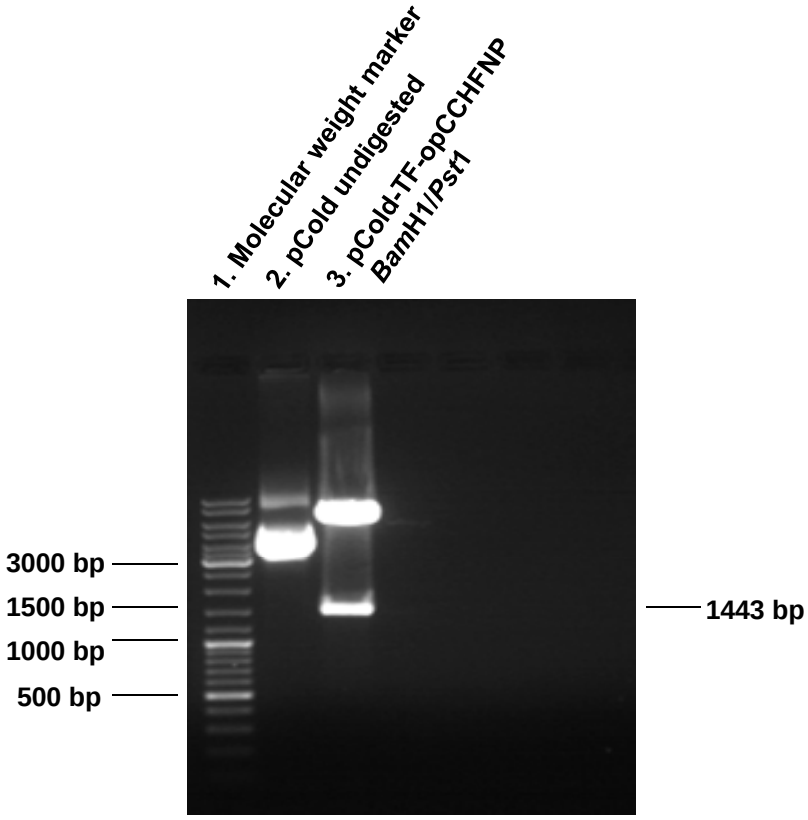


Figure 29. Agarose gel electrophoretic analysis of restriction enzyme analysis of a plasmid obtained from pCold TF vector. Lane 1 – O'GeneRuler DNA ladder mix molecular weight marker molecular weight marker; Lane 2 - 5µl of undigested pCold TF; Lane 3 - 20µl of pCold TF-opCCHFV NP digested with BamH1 and Pst1.

Verification of the presence of opCCHFV NP gene in pColdTF followed restriction enzyme analysis of pColdTF-opCCHFV NP plasmid with BamH1 and Pst1 yielding the predicted 1443 bp fragment as shown in Figure 29 (lane 3).

3.3.6 Sequencing and sequence analysis of the gene encoding CCHFV SPU415/85 in pCold TF vector

The nucleotide sequence of codon optimized gene encoding the NP from the first base of the pColdTF MCS is shown in Figure 30. Analysis of the sequence data confirmed insertional cloning of the correct codon optimized DNA fragment of 1443bp. The ORF was cloned with the BamH1 and Pst1 restriction sites as indicated in Figure 30 and is in frame with the start codon, MCS, TF tag, His tag sequence and stop codon of pColdTF vector.

Figure 30. Nucleotide sequence of the CCHFV codon optimized NP gene sequenced in pColdTF. The ATG start codon and TAG stop codon are highlighted in **grey.** The MCS of pColdTF is highlighted in **pink**. BamH1 and Pst1 restriction sites are highlighted in **red**.

CCHFV NP (pCold TF)

```
1                                                                    60
CATATGGAGCTCGGTACCCTCGAGGGATCCGAAAACAAAATCGAAGTGAACAACAAAGAT

61                                                                   120
GAAATGAACAAATGGTTCGAAGAATTCAAAAAAGGCAACGGTCTGGTTGATACCTTTACC

121                                                                  180
AATAGCTATAGCTTCTGCGAAAGCGTGCCGAACCTGGATAAATTTGTTTTCCAGATGGCG

181                                                                  240
AGCGCCACCGATGATGCACAGAAAGATAGCATCTATGCGAGCGCCCTGGTGGAAGCAACC

241                                                                  300
AAATTTTGTGCGCCGGTTTATGAATGCGCCTGGGTTAGCAGCACCGGCATTGTTAAACGT

301                                                                  360
GGTCTGGAATGGTTCGAAAAGAACGCAGGCACCATCAAAAGCTGGGATGAAAGCTATACC

361                                                                  420
GAACTGAAAGTGGATGTTCCGAAAATTGAACAGCTGGCGAACTATCAGCAGGCCGCACTG

421                                                                  480
AAATGGCGCAAAGATATCGGTTTTCGTGTGAATGCGAACACCGCCGCACTGAGCAATAAA

481                                                                  540
GTTCTGGCGGAATATAAAGTGCCGGGCGAAATTGTTATGAGCGTGAAAGAAATGCTGAGC

541                                                                  600
GATATGATCCGCCGTCGCAACCAGATTCTGAATCGTGGTGGCGATGAAAACCCGCGCGGT

601                                                                  660
CCGGTTAGCCGTGAACATGTGGATTGGTGTCGCGAATTCGTTAAAGGCAAATATATCATG

661                                                                  720
GCCTTTAATCCGCCGTGGGGTGATATTAACAAAAGCGGCCGTAGCGGTATCGCACTGGTG
```

721 780
GCGACCGGCCTGGCCAAACTGGCAGAAACCGAGGGTAAAGGCGTTTTCGATGAAGCAAAG

781 840
AAAACCGTGGAAGCCCTGAATGGTTATCTGGATAAACACAAAGATGAAGTTGATCGCGCA

841 900
AGCGCGGATAGCATGATTACCAACCTGCTGAAACATATCGCCAAAGCACAGGAACTGTAT

901 960
AAAAATAGCAGCGCGCTGCGTGCCCAGGGCGCACAGATTGATACCGCGTTTAGCAGCTAT

961 1020
TATTGGCTGTATAAAGCCGGTGTGACCCCGGAAACCTTCCCGACCGTTAGCCAGTTTCTG

1021 1080
TTCGAACTGGGCAAACAGCCGCGCGGTACCAAGAAAATGAAAAAAGCACTGCTGAGCACC

1081 1140
CCGATGAAATGGGGCAAAAAACTGTATGAACTGTTTGCGGATGATACCTTCCAGCAGAAC

1141 1200
CGTATCTATATGCACCCGGCCGTGCTGACCGCAGGTCGCATTAGCGAAATGGGCGTTTGC

1201 1260
TTTGGTACCATCCCGGTGGCGAATCCGGATGATGCCGCACAGGGCAGCGGTCATACCAAA

1261 1320
AGCATTCTGAACCTGCGTACCAATACCGAAACCAACAATCCGTGTGCCCGCACCATCGTT

1321 1380
AAACTGTTCGAAATTCAGAAAACCGGCTTTAACATCCAGGATATGGATATTGTGGCGAGC

1381 1440
GAACACCTGCTGCATCAGAGCCTGGTTGGTAAACAGAGCCCGTTTCAGAATGCGTATAAC

1441 1529
GTGAAAGGCAATGCGACCAGCGTGAACATCATTCTGCAGTCTAGATAG

3.3.7 Bacterial expression of pColdTF-opCCHFV NP using IPTG induction

The predicted size of the recombinant opCCHFV NP fused with trigger factor was 106kDa (54kDa NP plus 52kDa tag added by pColdTF). Bands indicating overexpressed proteins of sizes 106kDa and 52kDa (Trigger factor) are shown in Figure 31. The overexpressed recombinant opCCHFV NP fused to TF is demonstrated 4 and 24 hours post induction (lanes 4 and 6 repectively). The control plasmid expressing only TF protein is shown in lanes 5 and 7.

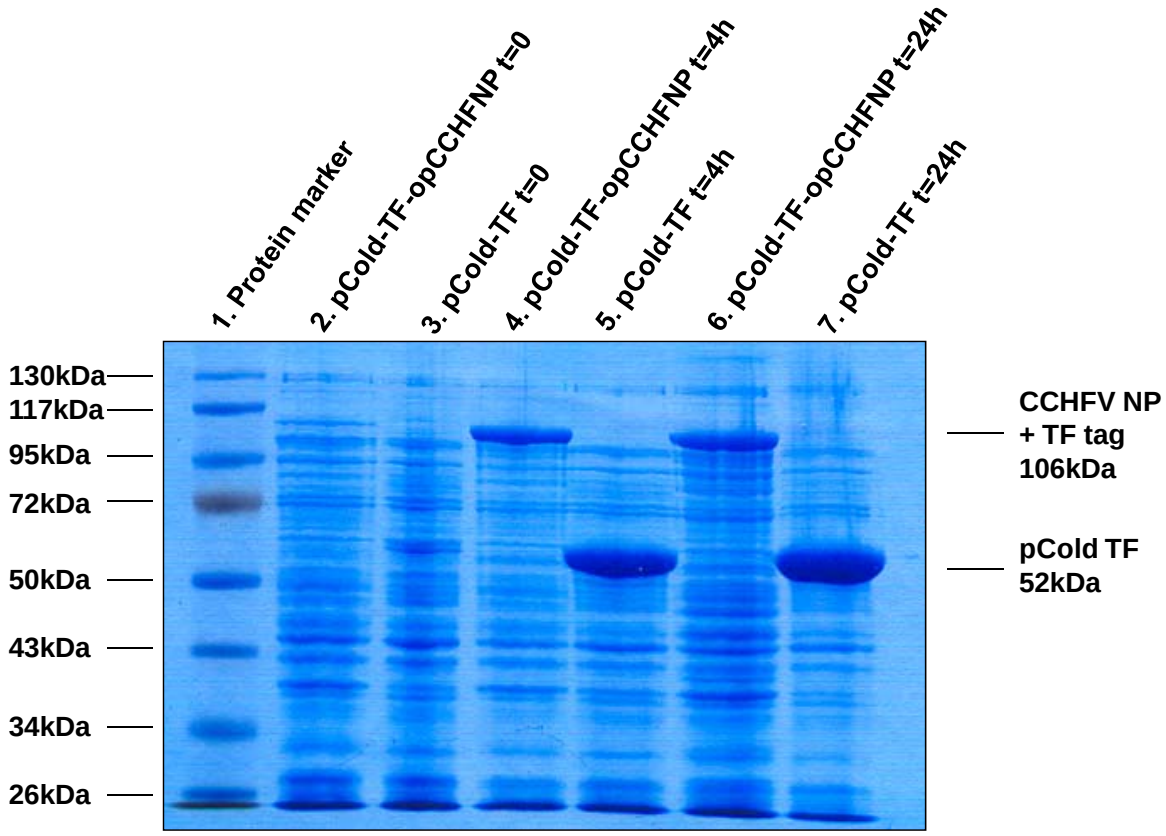


Figure 31. SDS PAGE analysis of proteins expressed using the recombinant expression vector, pColdTF-opCCHFV NP in OverExpress C43 (DE3) competent cells.

Lanes each contain 25µl of sample: Lane 1 – (Prestained Protein Ladder, Fermentas, USA); Lane 2 – pCold-TF-opCCHFV NP expression before induction; Lane 3 – pCold –TF negative control expression before induction; Lane 4 – pCold-TF-opCCHFV NP expression 4 hours after induction; Lane 5 – pCold –TF negative control expression 4 hours after induction; Lane 6 – pCold-TF-opCCHFV NP expression 24 hours after induction; Lane 7 – pCold –TF negative control expression 24 hours after induction.

3.3.8 Solubility, purification and characterization of His-tagged proteins

After performing a protein solubility study it was determined that the opCCHFV NP was expressed predominantly in the insoluble phase as shown in Figure 32. Lane 2 shows the total protein fraction expressed at 106kDa. In lane 3 an overexpressed protein representing the insoluble fraction is present at 106kDa and in lane4 a very faint protein band representing the soluble fraction is present. The band between 34kDa and 43kDa in the soluble and insoluble fractions was also present in the pColdTF fractions and is likely to be an *E. coli* protein. A similar result was obtained from fractions collected at t = 4. A number of attempts were made to increase the protein solubility by lowering the temperature at which the cells were growing from 16ºC to 10ºC during induction. However despite these attempts the protein still remained predominantly in the insoluble phase.

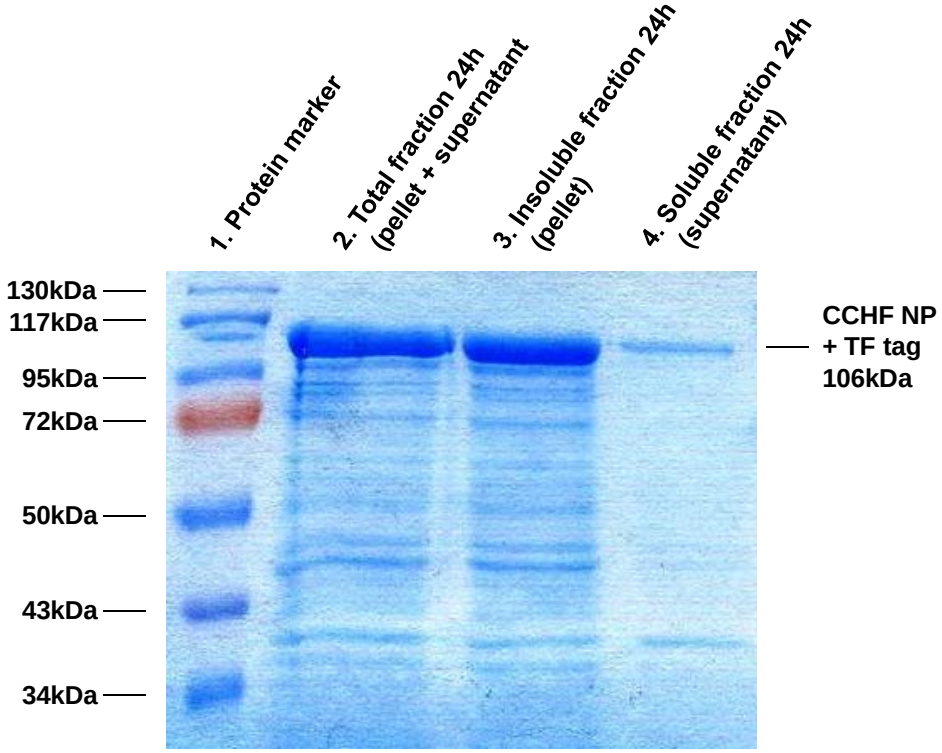


Figure 32. SDS-PAGE analysis of pColdTF-opCCHFV NP expressed protein fractions from solubility study performed 24h post induction with IPTG.

Lanes each contain 25µl of sample: Lane 1 – (Prestained Protein Ladder, Fermentas, USA); Lane 2 – Total fraction protein expressed 24 h after induction; Lane 3 – Insoluble fraction protein protein expressed 24 h after induction; Lane 4 – Soluble fraction protein expressed 24h after induction.

To harvest the protein from the insoluble phase, the pColdTF-opCCHFV NP and negative control pColdTF were purified using Protino nickel columns under denaturing conditions and the eluates collected from the column refolded to remove the urea and the protein analysed by SDS PAGE. A mock antigen prepared from pColdTF was purified under the same conditions. The purified proteins are shown in Figures 33 (control antigen) and 34 (CCHFV NP antigen). The protein concentrations measured were as follows: pCold TF (neg control)(Figure 36) = 1320.2 ng/µl; pColdTF-opCCHFV NP eluates 1 – 4 are 597ng/µl, 3221.5ng/µl, 329.2ng/µl and 157.9ng/µl respectively (Figure 34). Protein eluate 2 was selected and further characterization by Western blot to confirm the presence of a His tagged protein shown in Figure 35.

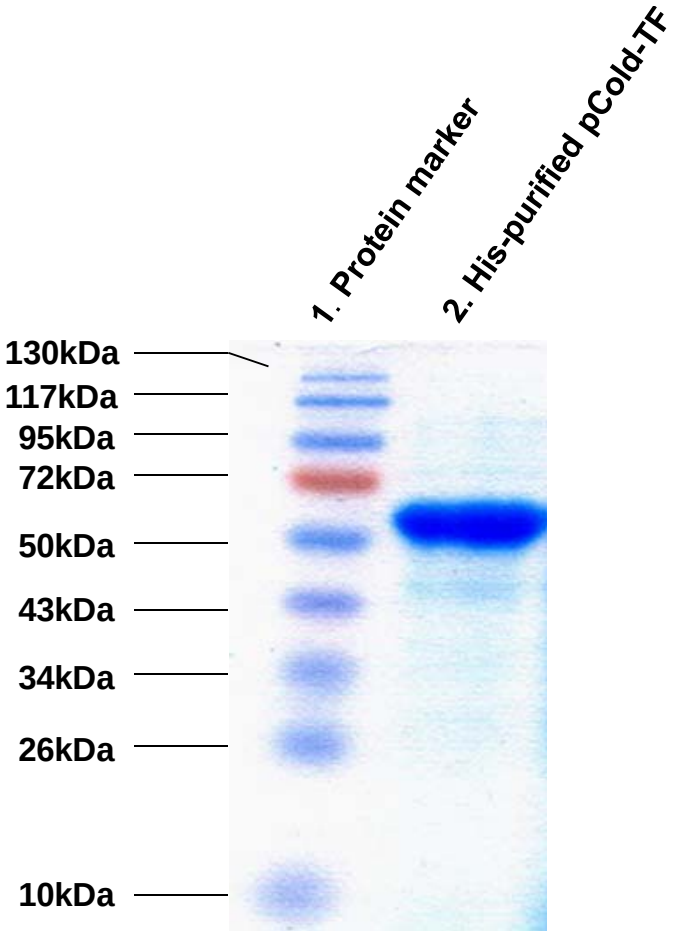


Figure 33. SDS PAGE analysis of His-purified pColdTF expressed as negative control 24 hours after induction.

Lane 1 – Prestained protein marker; Lane 2 - 25µl of 4.5ml of His-Tag purified pCold-TF

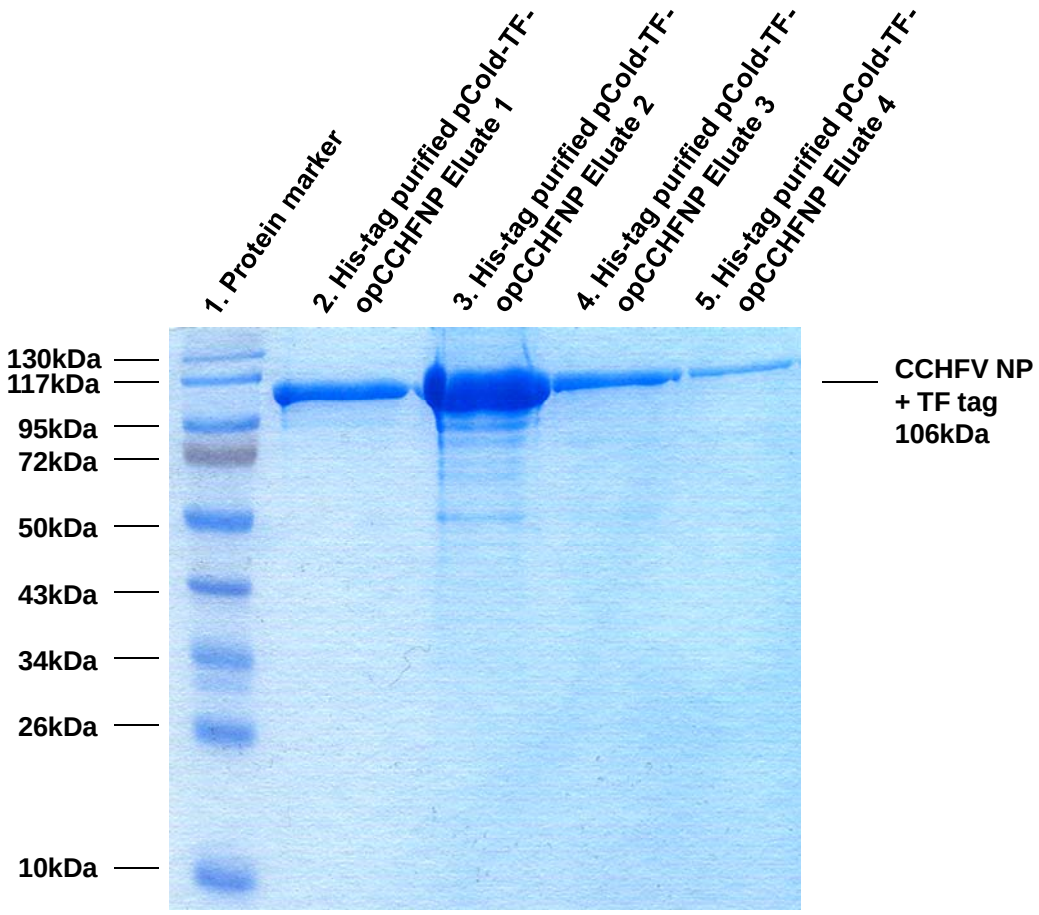


Figure 34. SDS PAGE analysis of His-purified pColdTF-opCCHFV NP eluates expressed as negative control 24 hours after induction.

Lanes contain 25µl of 4.5ml of His-Tag purified proteins. Lane 1 – Prestained protein marker; Lanes 2-5 - pCold-TFopCCHFV NP eluates collected from the 1000 packed Protino Ni-TED column

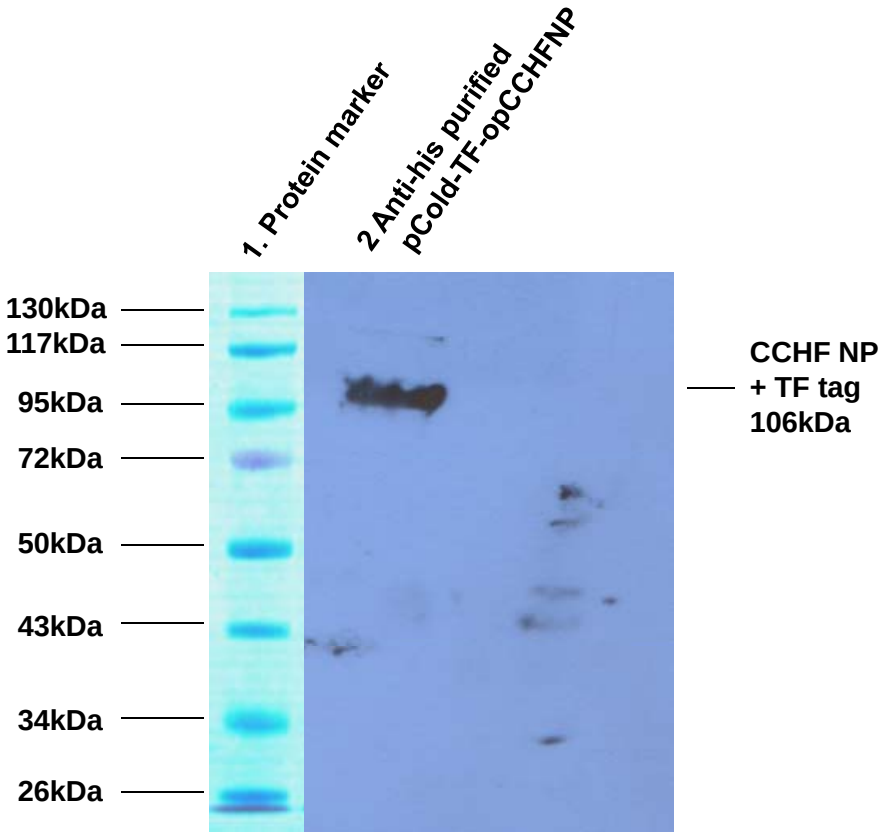


Figure 35. Western Blot analysis of recombinant protein expressed from His tagged pColdTF-opCCHFV NP

Lane 1 – Prestained protein marker; Lane 2- Anti-his purified recombinant NP.

From the scale-up of the cultures to produce large quantities of protein we calculated that a 250 ml culture would yield approximately 1.5 mg of protein. The protein was optimally diluted 1:2000 on the plates hence a culture of this volume would be sufficient to coat approximately 200 96-well ELISA plates.

3.4 Summary

Bacterial expression of recombinant proteins is generally preferred to mammalian and baculovirus expression systems. This can be attributed to the high-level expression of heterologous protein which can be produced at low costs and the large number of vectors and mutant strains readily available for use in bacterial expression systems (Jana and Deb, 2005).

The gene encoding the NP of CCHFV was cloned into two bacterial expression vectors namely pQE-80L and pCold TF. The pQE-80L vector utilizes high-level expression of 6x His-tagged proteins in *E.coli* at 37ºC and the target gene is under control of the T5 promoter, which is induced by IPTG. There was no evidence of detectable levels of CCHFV NP 1, 3 and 18 hours after induction of cultures using the construct with the native CCHFV NP gene. In contrast, a high yield of recombinant CCHFV NP was obtained from the pColdTF construct using the codon optimized gene. Altering experimental conditions such as temperature, incubation times and induction conditions failed to yield recombinant CCHFV NP at levels detectable on a protein gel.

Hence it was considered necessary to alter conditions to overcome potential codon bias. Robust expression of the recombinant NP was obtained after codon optimization. Overexpression however led to the formation of intracellular inclusion bodies. The insoluble to soluble ratio remained relatively unaffected by altering expression time or IPTG induction levels. Various methods are available to harvest proteins from inclusion bodies. The principle frequently involves denaturing the protein, purification of the protein utilizing the His tag and affinity to Ni^{2+} columns followed by renaturing and refolding of the protein. The protein was denatured with high concentrations of urea. The urea was subsequently removed by initially diluting the protein fraction 1:2 in LEW buffer to reduce the urea concentration to 4M and then removing the remaining urea by ultrafiltration over a period of 1 hour using a filter with a 30kDa cut off. The resultant protein was further characterized for use as a functional protein in bioassays.

CHAPTER 4

FUNCTIONAL CHARACTERIZATION OF BACTERIALLY EXPRESSED RECOMBINANT CCHFV NP USING A CODON OPTIMIZED GENE AND EVALUATION OF THE ANTIGEN FOR DETECTION OF ANTI-CCHFV IgG ANTIBODY

4.1 Introduction

A significant development in scientific research is the ability to clone a target gene and express a recombinant protein. Proteins that are pure, soluble and biologically active are in high demand in modern biotechnology (Sorensen and Mortensen, 2005). Bacterial host systems are probably the most frequently used to due simplicity in technology and cost effectiveness however the disadvantages are the lack of post-translantional modifications and the formation of inclusion bodies. The lack of post-translational modifications and production of inactive protein as a result of the formation of inclusion bodies poses a great challenge in bacterial expression systems (Sahdev et al., 2008).

Inclusion bodies are refractive intracellular particles of aggregated protein found in the cytoplasmic and periplasmic spaces of *E.coli* during high level expression of heterologous protein (Lilie et al., 1998; Fink, 1998; Singh and Panda, 2005). It is believed that high-level expression of proteins and highly hydrophobic proteins are more likely to lead to accumulation as inclusion bodies in *E.coli*. It is however possible to harvest protein from insoluble aggregates. There are two important issues to consider with regards to recovery of bioactive protein from inclusion bodies; solubilization of the protein aggregates usually with high concentrations of urea and subsequent refolding of the solubilized protein into a bioactive form. Refolding of proteins is an alteration in protein conformation from an unfolded to a folded or native state as schematically represented in Figure 36 (Tsumoto et al., 2003). Effectively refolding involves the removal of the denaturant at a suitable rate for refolding to occur (Tsumoto et al., 2003; Middelberg, 2004).

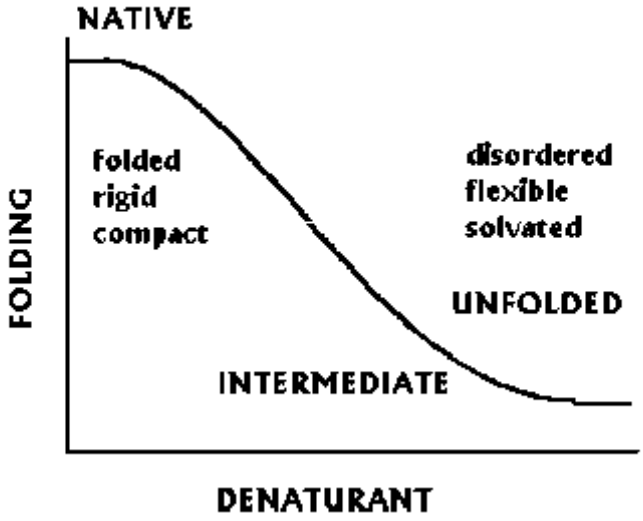


Figure 36. Conformation, flexibility and solubility of a protein as a function of denaturant concentration (Tsomoto et al., 2003). Degree of folding is plotted against increasing denaturant concentration. Physical properties of protein solution are given at high and low levels of concentration of denaturant.

Proteins are unfolded at high denaturant concentration whereas in aqueous buffer proteins are folded, firm and solid. The transfer of protein molecules from a high denaturant concentration to aqueous buffer should result in proper protein folding (Tsumoto et al., 2003; Singh and Panda, 2005). In the misfolded or aggregated state, where denaturants are not present, protein molecules lack flexibility to disaggregate and refold into the native form. As shown in Figure 36, a solution to refolding is in the intermediate concentration of the denaturant. The denaturant concentration should be low enough to induce protein collapse, but still allow the protein molecules to remain in solution and be flexible enough to refold. Equilibrium unfolding investigations have shown the presence of an intermediate (I) structure in the intermediate concentrations of the denaturant. These intermediate structures occur during refolding as demonstrated in Figure 37. The intermediate structure lacks stability and solubility and misfolds and aggregates as a result.

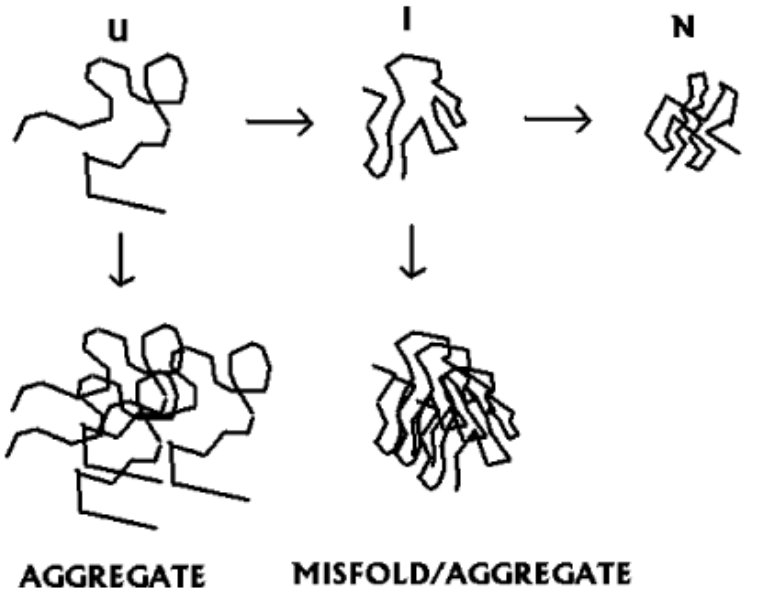


Figure 37. Schematic diagram of refolding course. U, I and N correspond to unfolded, intermediate and native state of protein respectively (Tsumoto et al., 2003).

Once the protein has been refolded it is necessary to determine if the protein is biologically active. For recombinant antigens part of the functionality of proteinsis to determine if any epitopes that may induce an immune response are adequately exposed. An epitope can be defined as a region on an antigen that interacts with antibodies thereafter inducing an immune response (Bui et al., 2007). For recombinant antigens that are to be used diagnostically their activity can be determined using immunoassays and known antibody positive serum samples.

The serological assays that have been used to diagnose CCHF infection include CF, HI, immunodiffussion, IFA and ELISA (Donets et al. 1982, Shepherd et al 1986, Shepherd et al. 1988). The use of IFA and ELISA have proven to be more reliable and useful in serological diagnosis as they are able to detect and differentiate IgM and IgG antibodies making recent and past infections easily distinguishable from one another (Shepherd et al. 1989a). Enzyme linked immunosorbent assays are fast, cost effective assays, readily automated with high sensitivity and specificity and are based on the ability of antibodies or antigens to bind to surfaces of ELISA plates. Ultimately these antibodies or antigens interact with proteins test samples forming antigen-antibody complexes. The interactions

are detected by a detector antibody labeled with an enzyme and the addition of a substrate which, in the presence of the enzyme, will undergo a colour change that can be detected colorimetrically. Evaluation of an in house developed ELISA requires adequate quality controls. Internal quality controls in the assay allow for continuous assessment of intra- and inter-plate variation. Information retrieved from these analyses can also be used to establish assay repeatability. Repeatability is defined as the ability of an assay to be performed on a daily basis and produce consistent results within a given range of the positive/negative controls (Jacobsen, 1998). Assay sensitivity, the ability of a test to detect the smallest positive result, and specificity, the ability of a test to detect the target for which it was specified must be considered.

In this chapter the biological functionality of recombinant CCHFV NP expressed in a bacterial system from a codon optimized gene, was assessed for use as a diagnostic tool. The protein was predominantly found in the insoluble phase as described in Chapter 3, hence was denatured and renatured during purification. After refolding of the harvested protein it was necessary to determine if epitopes required for antibody detection were exposed. Similarly to confirm the presence of immunogenic epitopes we aimed to induce a humoral immunity in mice that could be detected in immunoassays based on native CCHFV antigen. The usefulness of the recombinant protein as a diagnostic tool was investigated using convalescent patient sera. Whole viral antigens require BSL 4 facilities for preparation due to the risk to laboratory personnel and therefore involve high productions costs, hence it is important to prepare recombinant antigens as safe, low cost, sensitive tools for antibody detection. The ELISA procedure has been described in sections 4.2.5.1 and 4.2.5.2.

4.2 Materials and Methods

4.2.1 Human serum samples

A panel of 22 serum samples from 13 patients that have survived CCHFV infection were used in the study. The samples were screened for IgG antibody against the recombinant CCHFV NP and using a mouse brain derived antigen. Selected

samples were used to characterize the recombinant CCHFV NP using Western blot. Serum samples were collected from patients during the acute phase of illness and from 1 month up to 15 years after onset of illness. A panel of 14 negative serum samples collected from healthy patients with no history of CCHFV infection were used as negative controls to determine the cut off.

4.2.2 Western blot analysis for detection of anti-CCHFV IgG in human sera

The recombinant NP was characterized using Western blot analysis. After denaturation, the protein was purified on a Ni^{2+} column and each eluate fraction collected and ultrafiltered to refold the protein. The recombinant CCHFV NP fraction collected from the column (Chapter 3) with the highest protein concentration (3.2 µg/µl) was selected for characterization and used as an antigen in an ELISA. Western blot analysis of the protein was performed using the Pico Fast Western Blot Kit as previously described (MATERIALS AND METHODS, 3.2.14). Replicate aliquots of the recombinant NP were separated by SDS PAGE electrophoresis and transferred to nitrocellulose membrane. The recombinant antigen was reacted against each of the 16 selected human serum samples collected from convalescent patients with confirmed CCHFV infections. A serum sample collected from a volunteer with no known history of CCHFV infection and known to be antibody negative was used as the negative control. The membrane bound antigen was incubated with each serum sample diluted 1:100 in the fast western antibody diluent (supplied in the kit) for 30 min at 22-25ºC with shaking. The membranes were washed and processed as previously described (MATERIALS AND METHODS, 3.2.14), placed in a clear plastic wrap and exposed to film for approximately 60 sec. The film was incubated for 2 min in developer and for 1 min in fixer. The film was washed with water to remove excess fixer and visualized.

4.2.3 Immunization of mice with recombinant CCHFV NP protein

The purified recombinant NP protein was mixed with an aliquot of TiterMax® Gold adjuvant (Sigma-Aldrich, USA) as follows: 100µl protein (100µg) + 500µl adjuvant. The mixture was allowed to emulsify. A total of five eight week old female NIH

mice were used in this experiment. The mice were weighed prior to inoculations to determine their baseline weight (as required for animal ethics). The mice were immunized subcutaneously with a 100 µl inoculum of the recombinant NP in combination with adjuvant and monitored daily for any adverse events including irritation at the site of inoculation. The body weights of the mice were determined weekly to check for weight loss. Two weeks after inoculation the mice received identical booster immunizations. Mice were monitored daily and at six weeks post immunization they were anaesthetized using halothane and ex-sanguinated. All animal procedures were approved by the University of the Free State Animal Ethics Committee (ETOVS no 09/06).

TiterMax® Gold is an adjuvant that produces cell mediated and humoral responses in research animals. Adjuvants will promote the uptake of antigens by antigen presenting cells (APC), enhance stimulation of signals on APC required for helper T cell activation and subsequently proliferation of B cells. TiterMax® contains a block copolymer, CRL-8941, squalene, a metabolizable oil, and a stabilizer. The antigens are entrapped in a water-in-oil emulsion.

4.2.4 Detection of anti-CCHFV IgG antibody in immunized mice by IFA

The serum samples collected from immuinized mice were tested at the Special Pathogens Unit (SPU) at NICD/NHLS, Johannesburg for detection of IgG antibodies to CCHFV using IFA. Briefly, samples collected from the immunized mice were tested at doubling dilutions from 1:8 to 1:512 using antigen slides prepared from CCHFV infected cell cultures, and fluorescein-labelled (FITC) anti-immunoglobulin conjugate. Antigen slides were prepared at the SPU from Vero 76 cells infected with CCHFV. Infected cells were mixed with uninfected cells at a ratio of 1:10 and 10 µl aliquots of the cell suspension were applied to each well of a 8-well multi-test slide (Flow laboratories, UK), dried and fixed in cold acetone for a minimum of 20 min and stored at -70ºC until use. For the IF tests 10 µl aliquots of test sera, undiluted and diluted, were applied to each well of an antigen slide. The slides were incubated at 37ºC in a moist chamber for 20 min. The slides were washed in PBS for 3 min and 1 min in distilled water and dried. A 10 µl aliquot of FITC labeled anti human IgG conjugate (Zymed Laboratories, USA) diluted 1:100

in Evans blue as a counterstain was applied to each well and slides were incubated at 37ºC in a moist chamber for 20 min. The slides were washed in PBS for 3 min and 1 min in distilled water, dried and mounted with glycerol mounting fluid and coverslipped. The slides were read with a Nikon ultraviolet light microscope using a 20x fluorescent lens.

4.2.5 Enzyme-linked immunoassays

The ELISA were performed in 96 well immunoassay plates, and optimal working dilutions of the reagents were determined by checkerboard titrations. Throughout the assay, reagent volumes of 100 µl were used, the diluent for reagents was PBS containing 2% skimmed milk powder, incubations were performed for 1 hour at 37ºC, wells were blocked after coating with 200 µl PBS containing 10% skimmed milk powder and plates were washed thrice with PBS containing 0.1% Tween 20 unless specified. Optimal dilutions of the recombinant CCHFV NP antigen and pColdTF mock antigen were determined using serial fold dilutions from 1:500 to 1:4000.

4.2.5.1 IgG ELISA using recombinant antigen

Briefly, a 96 well microtiter Polysorb plate (Nunc Immunoplate, Denmark) was coated O/N at 4°C with recombinant CCHFV NP (3.2µg protein/µl) and mock pColdTF antigen diluted 1:2000 in PBS. The plate layout is illustrated in Figure 38. After the plates were washed and post-coated, human serum samples and controls, diluted 1:100 in diluent were added to each well in duplicate. The plates were incubated, washed and anti-human IgG horse radish peroxidase (HRPO) (Zymed laboratories, UK) diluted 1:4000 was added to each well. After further incubation and washing, positive reactors were visualized using the substrate 2,2'-Azino-di-ethyl-benzothiazoline-sulfonic acid peroxidase substrate (ABTS) (Kirkegaard and Perry Laboratories, USA). The plates were incubated at room temperature (22-25ºC) in the dark for 30 min and the OD values were read at 405nm. The net OD value for each test and control serum was determined.

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	1	2	2	3	3	4	4	5	5	C -	C -
B	6	6	7	7	8	8	9	9	10	10		
C	11	11	12	12	13	13	14	14	15	15		
D	16	16	17	17	18	18	19	19	20	20		
E	1	1	2	2	3	3	4	4	5	5	C -	C -
F	6	6	7	7	8	8	9	9	10	10		
G	11	11	12	12	13	13	14	14	15	15		
H	16	16	17	17	18	18	19	19	20	20		

CCHF NP Ag **Mock Ag** **C - : Negative control serum**

Patient 17 = Positive control serum

Figure 38. Plate layout for IgG ELISA using recombinant antigens.

4.2.5.2. IgG ELISA using whole virus mouse brain derived antigen

Serum samples were tested using an ELISA developed and validated for use in the SPU at NICD/NHLS. Briefly, monoclonal antibody against CCHFV NP diluted 1:2000 in PBS was used to coat a 96 well Maxisorb (Nunc Immunoplate, Denmark) microtitre plate O/N at 4ºC. After overnight coating the plates were washed and blocked with 10% skimmed milk/PBS. Sucrose acetone derived mouse brain CCHFV antigen or mock antigen diluted 1:400 was added to each well and the plates incubated for 1 hour. The plates were washed and serum samples were added to each well in duplicate at a dilution of 1:100 in 2% skimmed milk/PBS. The plates were further incubated and after washing anti-human IgG HRPO conjugate diluted 1:4000 was added. The plates were incubated as previously described, washed and positive reactors detected with ABTS substrate. The plates were reacted with ABTS for 30 min in the dark at 22-25ºC and the OD values were read at 405nm. Net OD values were determined.

4.2.6 Repeatability

To establish the assay repeatability, four serum samples selected randomly from the panel of convalescent patients were tested on a plate in duplicate on four separate occasions.

4.2.7. Selection of cut-off values

Cut off values to separate positive results from negative results were determined using a panel of 14 negative control sera. The ELISA was performed on 4 separate occasions using each serum sample to obtain values for 56 replicates. The mean net OD and standard deviation (SD) obtained from 56 replicates was calculated. A cut off value was determined from the mean net OD of the negative panel plus 2 SD values.

4.2.8 Determination of antibody titers against CCHFV in convalescent sera

The endpoint titer of a sample is defined as the reciprocal of the highest dilution that would result in a positive reaction in the assay (Frey et al., 1998). Therefore endpoint titers of our patient sera were determined to get a quantitative result. An IgG ELISA was performed as previously described (MATERIALS AND METHODS, 4.2.5.1) with a few modifications. The patient sera were first diluted 1:4 in PBS in suitable diluting plates. Following the blocking and washing step, 100µl of 2% diluent was added to each well. A 33µl aliquot of each serum sample was transferred from the diluting plate to the appropriate wells in the microtiter plates and subjected to 4-fold dilutions from 1:100 to 1:6400 as described in Figure 39. After diluting the sera at 1:6400 the remaining 33µl was discarded. OD values were read at 405nm after optimal development of substrate at 30 min and antibody titers were determined from net OD values.

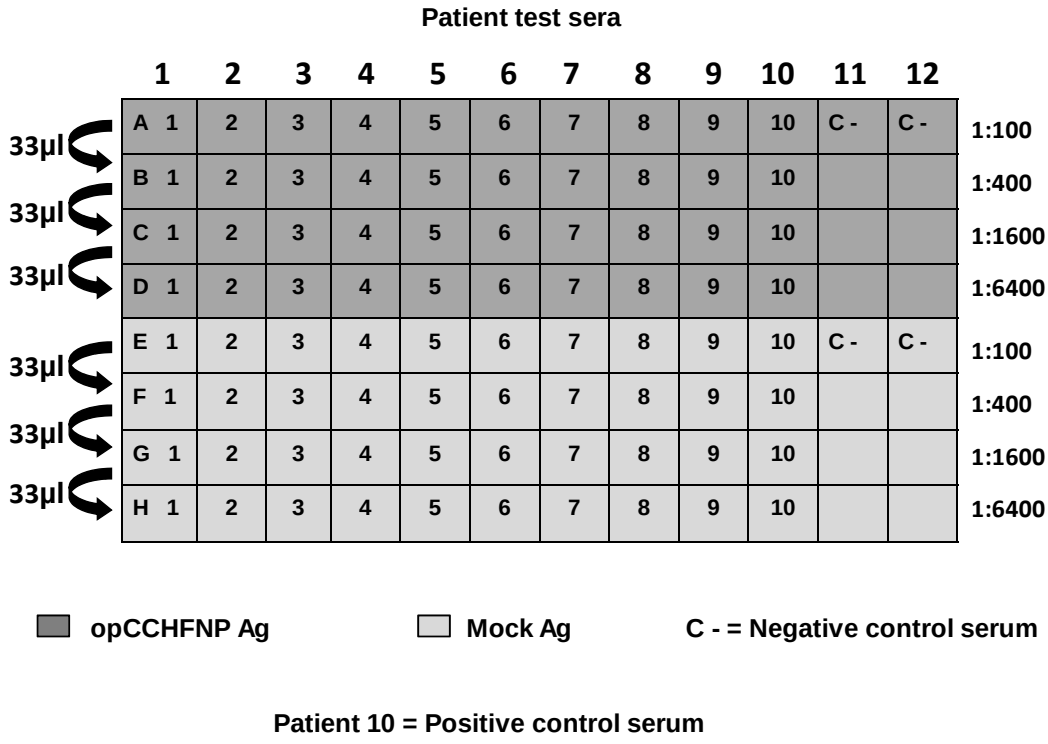


Figure 39. Plate layout for titration of convalescent patient sera in IgG ELISA.

4.2.9 Sensitivity of IgG ELISA using recombinant antigen

The sensitivity of the recombinant CCHFV NP antigen was compared with that of the mouse brain antigen for detection of IgG antibodies against CCHFV in the convalescent patient sera. A 2 × 2 contingency table was used to illustrate the sensitivity of the recombinant protein. Results in test sera that agree with the known positive and negative status of the sera in question were classified as true positive (TP) and true negative (TN), respectively. Results in test sera that do not agree with the known positive and known negative status of the sera in question were classified as false negative (FN) and false positive (FP), respectively (Simmons, 2008). The sensitivity (Sn) of the ELISA based on recombinant protein was determined using the formula:

$$\mathrm{Sn} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

4.2.10 Stability of bacterially expressed pColdTF-opCCHFV NP

A stability study was performed over a period of four months to test whether the recombinant CCHFV NP protein maintained biological activity over time. The recombinant NP and pColdTF mock antigens diluted 1:2000 were used to coat 5 PolySorb microtiter plates O/N at 4ºC. The following day the plates were washed and blocked in 10% skimmed milk as described (MATERIALS AND METHODS, 4.2.5.1). A known high positive serum sample was selected and used as the high positive control (C++), and a known low positive serum sample was selected and used as the low positive control (C+) to confirm theat the ELISA could detect samples with low antibody levels. A serum sample was selected from a healthy volunteer with no history of CCHFV infection and used as the negative control (C-). The plates were stored at 4ºC and tested at week 0, week 4, week 8, week 12 and week 16 after preparation respectively.

4.2.11. Statistical analysis of data

Normalization of data

OD values are absolute measurements that are influenced by variables such as temperature. To account for variability, results can be expressed as a function of the reactivity of control samples included in each run. Therefore absorbance or OD values were expressed as percentage positive (PP) relative to a high positive control serum. The following statistical calculations were used:

Positive to negative signal ratio:

Pos/Neg ratio = $OD_{C++/}OD_{C-}$

Net optical density (OD):

Net OD = OD in wells with virus antigen minus OD in wells with control antigen

Percent positivity (PP):

PP = (Mean net OD of test sample / mean net OD of C++) x100

C++ = high positive control

C - = negative control

4.3 Results

4.3.1 Induction of humoral antibody response using recombinant CCHFV NP

Five NIH mice were immunized with two doses of recombinant protein and adjuvant. TiterMax® Gold adjuvant has been developed to replace Freund's Complete Adjuvant and is safer and less toxic to the animals. The adjuvant should stimulate both cell mediated and humoral antibody responses in research animals. At five weeks post immunization blood samples were collected from euthanased mice. Three mice developed a detectable antibody titer ranging from 1:16 to 1:128, one mouse had a positive reaction using undiluted sera and one mouse had no detectable antibody by IFA against CCHFV.

4.3.2 Detection anti-CCHFV IgG antibody in human sera

4.3.2.1. Western blot analysis

Anti-CCHFV IgG was detected in 13/16 human serum samples tested by Western blot analysis as shown in Table 13 and selected serum samples illustrated in Figure 40, Acute phase samples collected on days 14 to 16 did not react on Western blot.

Table 13: Serum samples collected from 13 patients at various days after onset of illness. Detection of anti-CCHFV IgG in patient serum samples was performed using Western blot, ELISA with recombinant NP and ELISA with mouse brain derived antigen.

Laboratory number	Patient No	Duration after illness	IgG ELISA rec antigen (antibody titer)	Western Blot	IgG ELISA mouse brain derived antigen
VBD 6/08	1	8 years	Pos (1:400)	Pos	N/A
VBD 71/08	1	8 years	Pos (1:1600)	Pos	Pos
VBD 53/10	1	10 years	Pos (1:1600)	N/A	N/A
VBD 10/08	2	15 days	Pos (1:100)	Neg	Pos
VBD 11/08	2	16 days	Pos (1:100)	Neg	Pos
VBD 70/08	2	10 months	Pos (1:400)	Pos	Pos
VBD 52/10	2	2 years	Pos (1:400)	N/A	N/A
VBD 22/08	3	14 days	Pos (1:100)	Neg	Pos
VBD 66/08	3	9 months	Pos (1:100)	Pos	Pos
VBD 42/10	3	2 years	Pos (1:100)	Pos	Pos
VBD 67/08	4	7 years	Pos (1:400)	Pos	Pos
VBD 41/10	4	2 years	Pos (1:400)	Pos	Pos
VBD 69/08	5	4 years	Pos (1:400)	Pos	Pos
VBD 6/09	6	4 months	Pos (1:1600)	Pos	Pos
VBD 10/09	7	2-3months	Pos (1:100)	N/A	Pos
VBD 57/09	8	11 years	Pos (1:100)	Pos	Pos
VBD 53/09	9	9 years	Pos (N/A)	Pos	Pos
VBD 30/10	9	10 years	Pos (1:1600)	Pos	Pos
VBD 29/10	10	19 days	Pos (1:1600)	Pos	Pos
VBD 51/10	11	23 months	Pos (1:100)	N/A	N/A
VBD 6/11	12	11 months	Pos (1:100)	N/A	N/A
VBD 29/11	13	10 months	Pos (N/A)	N/A	N/A

Pos: positive

N/A :not available

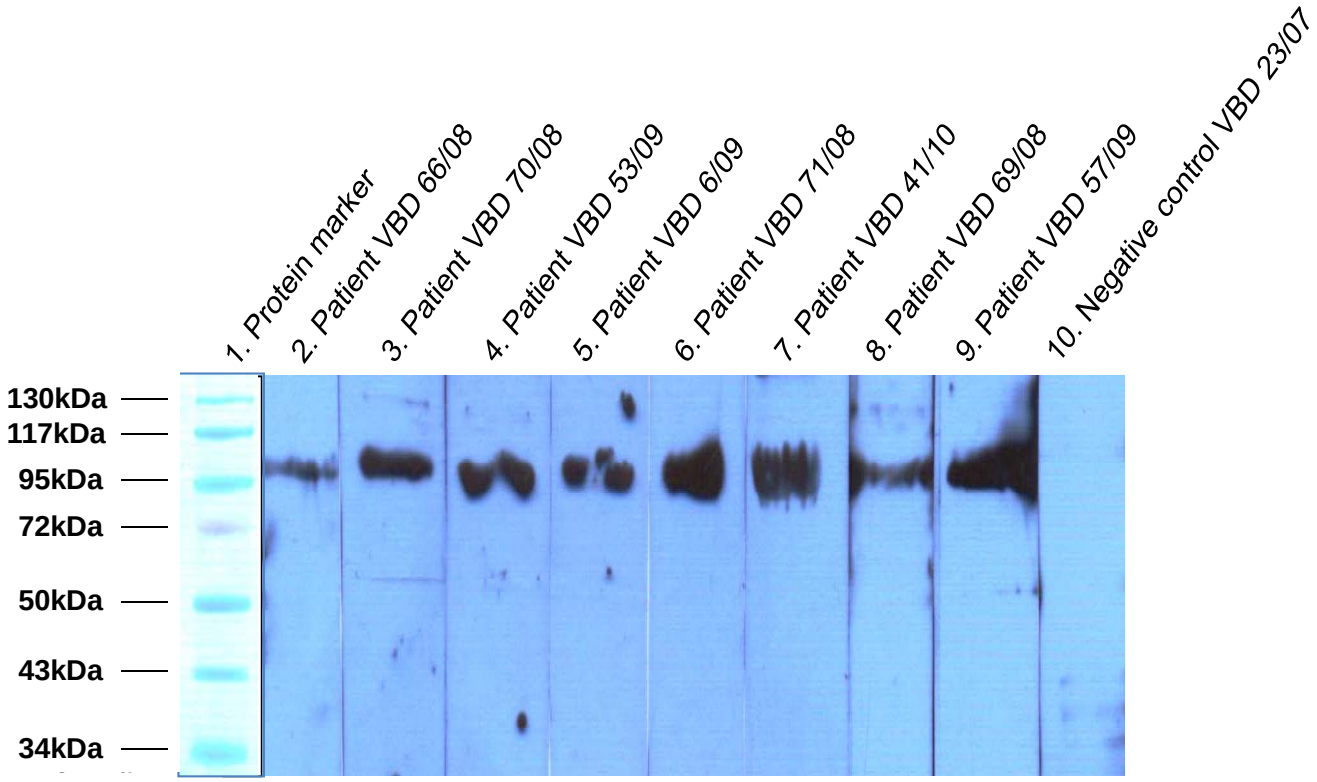


Figure 40. Anti-CCHFV IgG in human sera reacted with recombinant NP in a Western blot

Lanes each contain 25µl of sample: Lane 1 – (Prestained Protein Ladder, Fermentas, USA); Lane 2 – VBD 66/08; Lane 3 – VBD 70/08; Lane 4 – VBD 53/09; Lane 5 – VBD 6/09; Lane 6 – VBD 71/08; Lane 7 – VBD 41/10; Lane 8 – VBD 69/08; Lane 9 – VBD 57/09; Lane 10 – Negative control VBD 23/07.

4.3.2.2 IgG ELISA using recombinant antigen

The result of the checkerboard titration using two-fold dilutions of antigen from 1:500 to 1:4000 is illustrated in Figure 41. The level of positive to negative signal ratio was established by a specific ELISA using controls is an indication of the performance of the ELISA. A ratio greater than 1:10 indicates that the ELISA can detect and differentiate between sera of various antibody titer. The line representing the different dilutions peaked at the antigen dilution between 3 and 3.5 $\log_{10}$. In subsequent ELISA the recombinant NP antigen was diluted 1:2000.

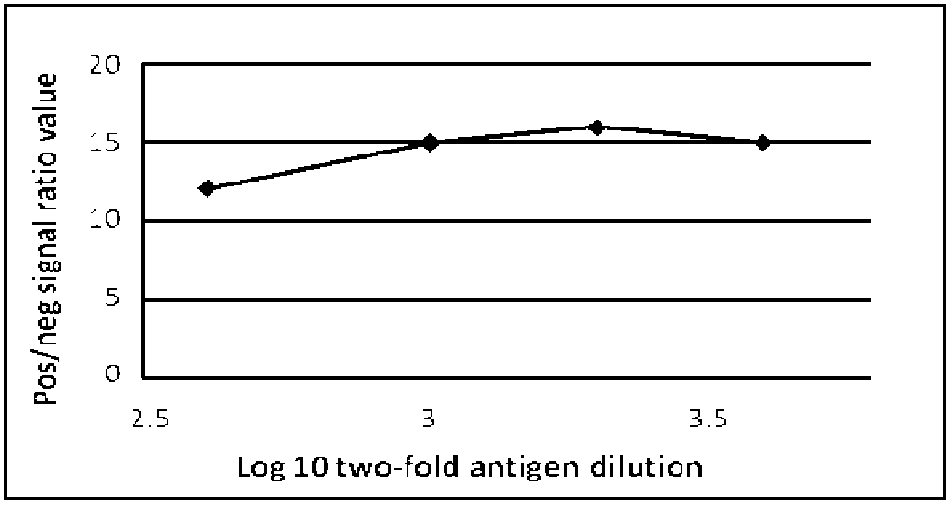


Figure 41. Positive to negative signal ratios of the indirect ELISA

As shown in Table 13, anti-CCHFV IgG antibody was detected in 16/16 serum samples from survivors of CCHFV infection. The antibody titers of the patient sera are also shown in Table 13. The patients had antibody titers ranging from 1:100 to 1:1600.

A total of 22 human sera were reacted against the recombinant NP in an indirect ELISA. Negative serum samples were used to determine a cut off and to establish a level of nonspecific background. Similarly the human sera were reacted against mock antigen to determine background levels. Serum samples were collected from acute and convalescent patients from day 14 after onset of illness up to 15 years. An IgG antibody response was detected in all the samples tested using the ELISA with recombinant NP. There was 100% concordance with the test results obtained from sera tested using a mouse brain derived antigen (see Section 4.3.2.3.). In contrast the Western blot was less sensitive for detecting IgG antibody and was not able to detect antibody in serum samples collected during the acute stage, days 14 to 15.

4.3.2.2.1 Selection of cut off

The high and low ranges of the OD and PP values of the negative panel used to determine the cut off are shown in Table 14. The cut off could be expressed as either OD or PP value determined from the mean net OD and mean net PP value

of the negative panel plus 2 SD are shown in Table 15.

Table 14. High and low ranges of the OD and PP values of negative panel used to determine cut off.

Laboratory Number	OD[a] ranges		PP[b] ranges	
	High	Low	%High	%Low
VBD 17/07	0.138	0.064	11.794	5.470
VBD 20/07	0.182	0.060	15.555	5.128
VBD 21/07	0.057	0	4.871	0
VBD 22/07	0.029	0.010	2.478	0.850
VBD 23/07	0.198	0.090	16.923	7.692
VBD 25/07	0.193	0.074	16.495	6.324
VBD18/08	0.142	0.044	12.136	3.760
VBD19/08	0.137	0.039	11.709	3.333
VBD 28/08	0.154	0.043	13.162	3.675
VBD 43/09	0.223	0.033	19.059	2.820
VBD 44/09	0.24	0.046	20.512	3.931
VBD 47/09	0.181	0.047	15.470	4.017
VBD 48/09	0.102	0.067	8.717	5.726
VBD 50/09	0.237	0.069	20.256	5.897

[a]OD = Optical density
[b]PP = Percent positivity

Table 15. Cut off value determined from negative panel of sera.

Serum	[a]*n*	[b]Mean	[c]SD	Cut off	
				[d]OD	[e]PP
Negative panel	56	0.0953	0.0721	0.240	20.512%

[a]*n*= number of replicates
[b]Mean = mean NET OD of replicates
[c]SD = standard deviation
[d]OD = optical density
[e]PP = percent positivity of high positive control serum

4.3.2.2.2. Stability of bacterially expressed recombinant NP

The PP values obtained for C++ serum decreased over the testing period of 4 months that the plates were stored. The PP values for C++, C+ and C- sera are shown in

Figure 42. The PP values of the high positive (C++) and low positive serum (C+) were approximately 140% and 50% respectively immediately after antigen preparation. Two weeks later the protein remained relatively stable as the PP values of C++ and C+ decreased by approximately 5 and 10% respectively. Although the values decresed with time the high and low positive sera were still detected at levels above the cutoff of 20.512% 16 weeks after preparation. (RESULTS 4.3.2.2.1)

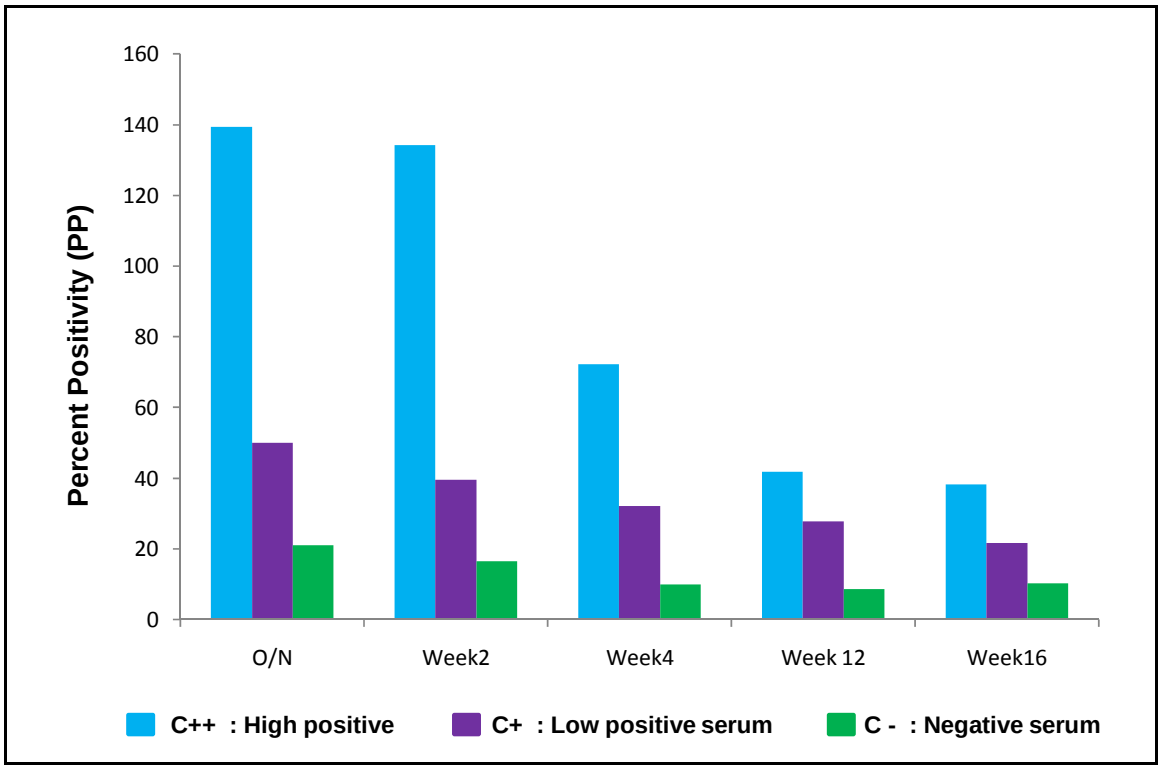


Figure 42. Analysis of recombinant CCHFV NP stability over 4 months.

4.3.2.2.3. Interassay variability of recombinant antigen ELISA

To determine the inter assay variability four samples were tested on different days and the data normalized by calculating the ratio of the net OD value obtained for each serum sample and the net OD value obtained for a negative control. Similar OD values were obtained as illustrated in Figure 43.

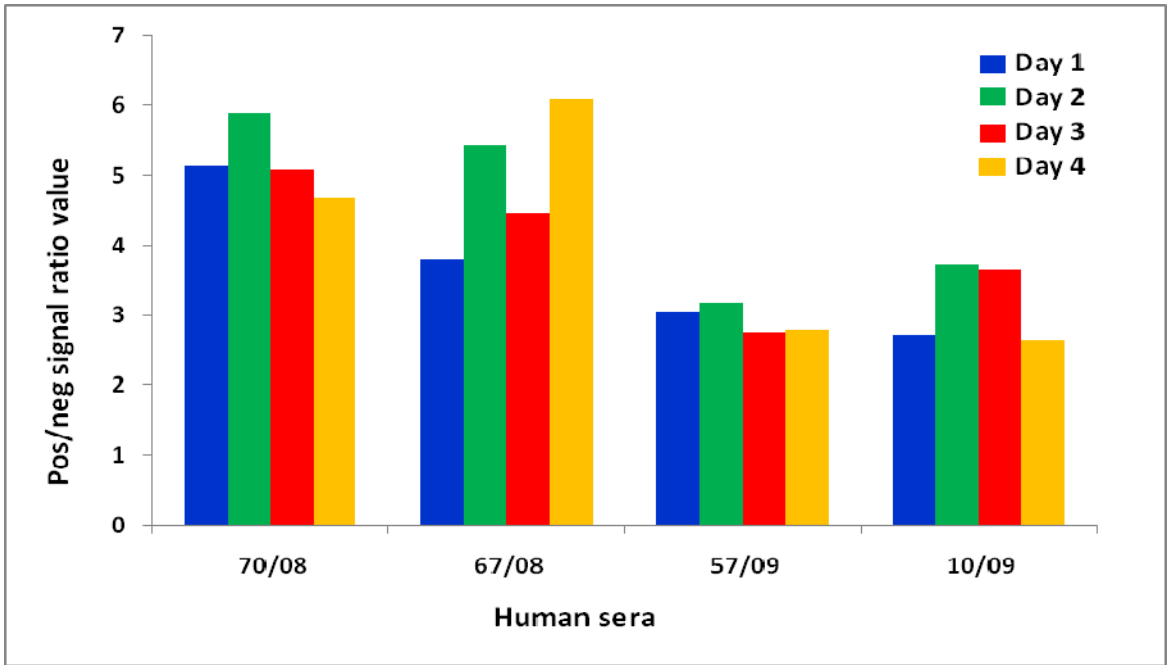


Figure 43. Positive to negative signal ratios of net OD values obtained for four human serum samples from survivors of CCHFV infection (VBD 70/08, 67/08, 57/09 and 10/09).

4.3.2.2.4. Detection of IgG antibodies in human sera using ELISA

The net OD values (included in Appendix 1) were used to calculate PP values shown in Table 16. Figure 44 illustrates the mean PP values and range obtained for each serum sample tested in duplicate and the mean PP values and range for each negative serum sample tested over four consecutive runs. The cut off expressed as a PP value is 20.95%. Serum samples with Lab number 10/08, 11/08 and 22/08 were collected from patients during the acute phase of illness and did not react in the Western blot (RESULTS, 4.3.1) but were detected by the ELISA. This suggests that the ELISA was more sensitive in detecting acute phase patients than Western blot.

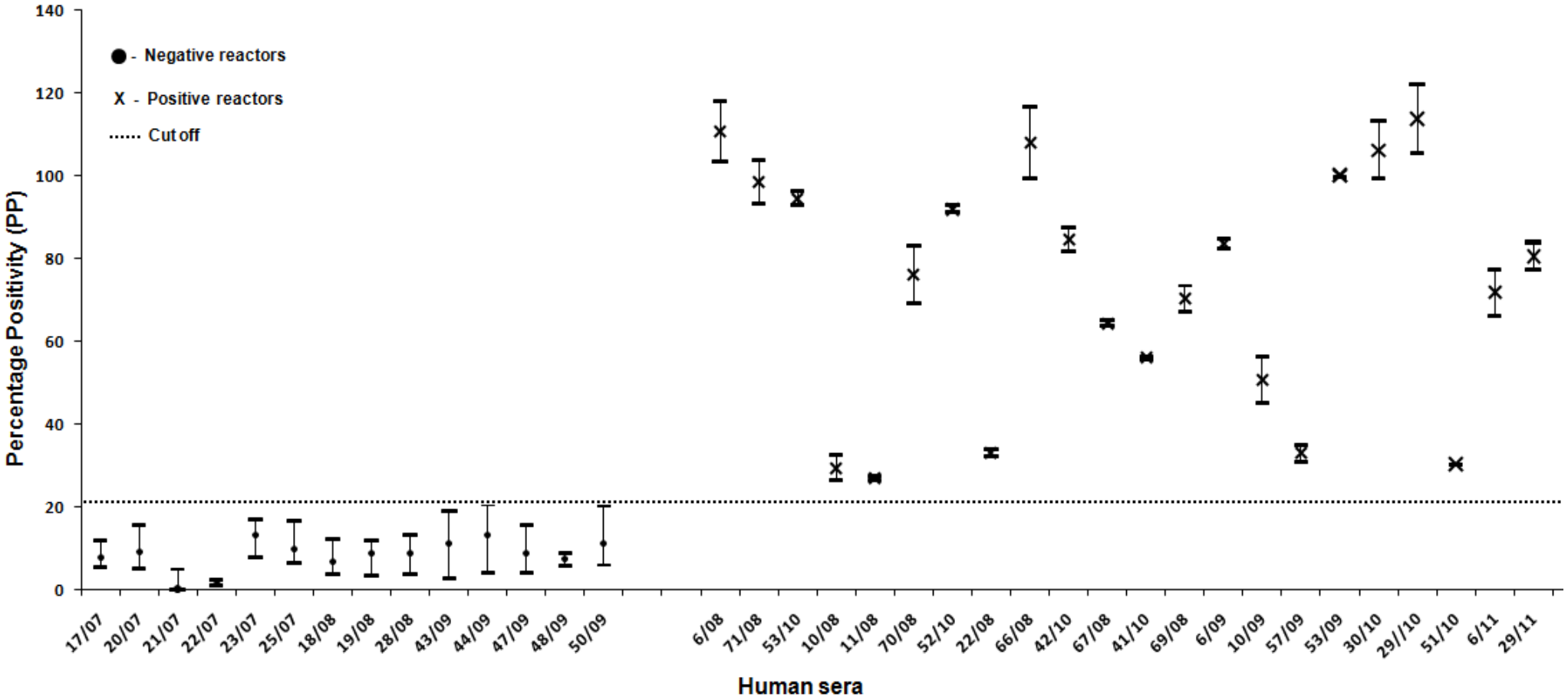
Table 16. Average net OD and PP values of convalescent patient sera calculated using high positive control.

Laboratory number	Patient No	Average NET [a]OD	Mean net OD of C++ = 1.17 Average % [b]PP
VBD 6/08	1	1.296	110.763
VBD 71/08	1	1.153	98.545
VBD 53/10	1	1.144	94.575
VBD 10/08	2	0.344	29.45
VBD 11/08	2	0.314	26.884
VBD 70/08	2	0.890	76.194
VBD 52/10	2	1.120	92
VBD 22/08	3	0.399	32.998
VBD 66/08	3	1.263	107.943
VBD 42/10	3	0.989	84.571
VBD 67/08	4	0.753	64.412
VBD 41/10	4	0.655	56.023
VBD 69/08	5	0.823	70.293
VBD 6/09	6	0.978	83.584
VBD 10/09	7	0.594	50.766
VBD 57/09	8	0.385	32.905
VBD 53/09	9	1.170	100
VBD 30/10	9	1.050	106.236
VBD 29/10	10	1.295	113.806
VBD 51/10	11	0.354	30.298
VBD 6/11	12	0.840	71.799
VBD 29/11	13	0.943	80.597

[a]PP= percent positivity

[b]OD = Optical density

Figure 44. Detection of anti-CCHFV IgG by ELISA using recombinant NP. The average PP values (+/- range) of the positive patients are indicated with an **X** symbol and the average PP values (+/- range) of the negative panel tested over four days is indicated with a ● symbol.



4.3.2.2.5. Sensitivity of the recombinant NP antigen

Results from the ELISA performed using mouse derived antigen were available for 16/22 serum samples included in the study. The 16 convalescent patient samples that tested positive for CCHFV antibodies by ELISA using the mouse brain derived antigen also tested positive for CCHFV antibodies using the recombinant antigen as shown in Figure 45 and Table 13. Our results showed 100% concordance with the results obtained from an ELISA using mouse brain derived antigen (Table 13 and Figure 45).

		SPU IgG ELISA	
		Positive	**Negative**
Recombinant IgG ELISA	**Positive**	**16** **TP**	**0** **FP**
	Negative	**0** **FN**	**0** **TN**

Figure 45. A 2 x 2 contingency table illustrating assessment of the recombinant NP.

4.4 Summary

This chapter describes experiments performed to determine the biological functionality of the bacterially expressed CCHFV NP from a codon optimized gene. Since the protein was predominantly found in the insoluble phase as described in Chapter 3, it was necessary to denature and refold the protein during purification. Subsequent to refolding the protein it was important to determine if the protein was able to react against human serum samples using Western blot analysis and if there were epitopes exposed that could induce humoral immunity in mice that could be detected using IFA.

Patients that survive CCHFV infections have a demonstrable IgG and IgM antibody response from days 5 to 7 after onset of illness (Shepherd et al., 1988, Burt et al., 1994). Antibody detection is usually performed by ELISA or immunofluorescence assays. Recombinant protein technology would be a useful and cost effective platform for preparation of safe antigens. As this work aimed to use the NP for developing assays for diagnostic purposes and for performing serological surveillance, various studies were performed to confirm that the fusion protein expressed from the codon optimized and synthesized gene was functionally active and that antibodies in human convalescent sera were reactive with the epitopes exposed on the bacterially-expressed NP. The protein was reacted against serum samples collected from patients at varying stages after onset of illness from acute to convalescent stages using both an ELISA and a Western blot. The recombinant antigen was able to detect IgG antibody in all sera using ELISA and in convalescent sera using the Western blot technique. To confirm that the protein was able to induce a humoral antibody response that could be detected using CCHFV antigen derived from live virus, we immunized mice and tested serum samples using IF slides prepared from CCHFV infected Vero cells. In addition we were able to induce a detectable IgG antibody response in mice immunized using the NP. Hence the results suggest that bacterially expressed proteins lacking post translational modifications and folding that occur with mammalian and baculovirus expression systems have potential for use in both diagnosis and surveillance.

CHAPTER 5

DISCUSSION

CCHFV is an enveloped single stranded negative sense RNA virus belonging to the *Nairovirus* genus of the *Bunyaviridae* family. The genome comprises a S (small), M (medium) and L (large) segment which code for the nucleocapsid, the envelope glycoproteins G_n and G_c, and an RNA dependent RNA polymerase (Clerx et al. 1981; Elliot et al.1990; Marriott and Nuttall, 1992).

In the past ten years CCHFV has emerged in several countries in the Balkans and re-emergence in south-western regions of the Russian Federation in 1999 following 27 years without human cases of disease (Maltezou et al., 2010). Factors such as climate, changes in land and agricultural use, hunting, import and export of livestock contribute to the emergence of CCHFV as they influence ticks and their hosts and subsequently the epidemiology of CCHFV. Ticks belonging to the genus *Hyalomma* are present in southern and south-eastern Europe which raises concerns that the virus could spread to currently non-endemic areas in Europe (Ahmed et al., 2009; Maltezou et al., 2010). These concerns highlight the importance of developing standardized reagents and increasing diagnostic and surveillance capacity. Future research in the field of CCHFV will require the development of safe, standardized, low cost reagents and assays for diagnosis and surveillance, new methods for drug development, vaccine production and methods of vector control (Ahmed et al., 2009; Maltezou et al., 2010). Diagnosis of CCHFV infections during the acute phase is based on isolation of the virus or amplification of viral RNA. Patients that survive the infection have a demonstrable IgG and IgM antibody response usually from day 5 to 7 after onset of illness (Shepherd et al., 1988; Burt et al., 1994). Current serological diagnostic assays based on ELISA or IF use inactivated virus which requires BSL4 facilities for culturing the virus and therefore limits the number of laboratories that can prepare suitable reagents. Preparation of recombinant antigens would enable laboratories to perform serological diagnosis of CCHFV infections and surveillance studies, particularly in countries with less sophisticated facilities.

The aims of this study were to prepare a recombinant CCHFV nucleoprotein using a bacterial expression system, to determine if the protein was immunogenic and to determine if the protein was able to detect IgG antibodies in survivors of CCHF infection.

Sequence data from complete S segments available on GenBank were aligned and used to identify primers that would amplify the gene encoding the NP by RT-PCR. The identified primers were modified to include BamH1 and Pst1 restriction sites at the 5' ends to facilitate cloning into bacterial expression vectors. The amplicon was purified and cloned into pGEM® T Easy vector using T/A cloning. Constructs were confirmed using restriction enzyme analysis and DNA sequencing. The recombinant CCHFV NP plasmid was sequenced to determine the nucleotide sequence of the SPU415/85 NP gene which was subsequently aligned with sequence data from GenBank to confirm that the correct gene had been amplified and cloned into the vector for downstream cloning and expression applications.

Initially we aimed to express the native gene using a bacterial expression system and the NP gene was rescued from the recombinant plasmid and cloned into pQE-80L vector using the BamH1 and Pst1 restriction sites engineered into the RT-PCR primers and present in the multiple cloning site on the vector. Positive transformants were confirmed by restriction enzyme digestion and sequencing. Sequence data was analysed to confirm that the correct gene had been inserted and was in-frame with the start codon of the expression system and the 5' HIS tag. The vector is supplied with stop codons in each of the three coding frames. Various attempts were made to express the CCHFV NP protein. However no protein was detectable using SDS PAGE methods or Western blot. It is difficult to predict the reason why proteins are not expressed. Aside from hydrophobicity, proteins may be toxic to the cells during the growth phase, particularly at higher temperatures, or there may be insufficient IPTG in the samples to induce expression. Optimization was performed using a range of high and low IPTG concentrations with lower incubation temperatures to reduce the growth rate yet induce expression of protein that may have been present in very low concentrations.

Given that there are no conclusive rules that can be used to predict whether a protein will be expressed from a heterologous host it is possible to optimize a gene so that the codons are modified to take into account the codon bias of a particular host. Therefore the nucleotide sequence that we had determined for the open reading frame of our gene encoding the NP was analysed using the Rare Codon Analysis Tool software (download available on GenScript website at http://www.genscript.com). The CAI is a measure of codon bias and indicates the similarity between the codon usage of a gene and the codon usage of the host. The CAI is measured from 0 to 1 with 1 indicating no codon bias. Analysis of our sequence data for the CCHFV isolate that we had selected gave a CAI value of 0.64 and analysis of the frequency at which each codon on our gene was utilized by the *E.coli* host ranged from 10% to 70% along the length of the gene. Hence we elected to codon optimize the gene for expression in *E. coli*. After optimization, the codon usage increased with values ranging from approximately 70% to 100%, and the CAI of the codon optimized gene was estimated to be 0.95. The optimized gene was synthesized by GenScript and supplied cloned in the multiple cloning site of pUC57 with BamH1 and Pst1 restriction site modifications at the 5' end and 3' end respectively. The optimized gene encoding the nucleoprotein of CCHFV NP was excised from pUC57 using the BamH1 and Pst1 sites and cloned into pColdTF bacterial expression vector. We selected to use a different expression system which expresses at a lower temperature and uses a chaperone protein to assist with folding and solubility of expressed proteins. Previous attempts to express the native gene from pColdTF were not successful (data not shown). Cloning of the optimized gene was confirmed by restriction site analysis and sequencing. A 106 kDa protein was expressed from the construct likely representing the HIS tagged TF chaperone protein fused to the CCHFV NP protein. Western blot analysis confirmed the presence of a HIS tagged protein of 106kDa. A higher yield of the protein was present in the insoluble phase and optimization of the growth and induction conditions did not significantly alter the insoluble to soluble ratio of the expressed protein. Hence we attempted to harvest the protein from the insoluble phase by denaturing, purification and subsequent refolding of the protein.

Recombinant proteins must retain their biological functionality, and epitopes inducing humoral immune responses need to be exposed as they would be in the native protein, for the recombinant proteins to be of use in immunoassays. Bacterial systems lack many post translational modifications and the process of denaturing and refolding the protein introduce steps that could significantly influence the folding of the recombinant protein and reduce its biological activity. Hence it was important to not only confirm the biological activity of the recombinant protein using immunoassays but we also opted to immunize mice and determine if the antibodies induced by our recombinant protein could be detected using an antigen prepared from whole virus. The renatured protein, in combination with TiterMax Gold adjuvant, was used to immunize mice and 4/5 mice developed a detectable IgG antibody response to CCHFV NP using an IF assay. The lack of a detectable antibody response in one animal is likely a consequence of inadequate dose. We were satisfied that our recombinant antigen was able to induce a humoral antibody response recognized by whole viral antigen slides

Serum samples from acute and convalescent patients collected at varying stages after onset of illness were reacted in a Western blot with the recombinant CCHFV NP protein. The recombinant antigen was able to detect IgG antibody in all the convalescent patient sera but two serum samples collected on days 14 and 15 during the acute phase were not detected. In contrast all the samples were detected using the recombinant antigen in an ELISA. This could be a reflection of the sensitivity of the Western blot or that different epitopes are dominant during the acute phase. However as our final application for the NP is a diagnostic and surveillance tool the ELISA results was of more significance.

Previous workers have used recombinant antigens to detect IgG antibody against CCHFV. CCHFV NP expressed as a fusion protein in a baculovirus system was used to detect IgG antibody in serum samples from nine patients. Positive IgG results were only obtained for two patients in samples collected on days 5 and 11 (Saijo et al., 2005). The antigen was subsequently used prospectively on one patient and could detect an IgG response in sera collected on days 5 and 9 (Tang et al., 2003). A recombinant NP expressed in mammalian cells via a Semliki Forest alphavirus

replicon system was used in IFA tests and ELISA and was able to detect IgG antibody in 10/29 sera collected within eight days after onset of illness from patients with suspected CCHFV infections (Garcia et al., 2006). Hence the current data available for use of recombinant antigens to detect CCHFV antibody responses is very limited. Bacterially expressed proteins have not previously been described for detection of antibody in clinical samples. Bacterial expression would be the most cost effective and technically less challenging technique.

We tested 22 serum samples from patients that had confirmed infections. Due to the biohazardous nature of samples collected during the acute stage of illness we only tested samples collected two weeks after onset of illness which could be safely handled in our laboratory with BSL 2 facilities. Our results showed 100% concordance with the results obtained from an ELISA using mouse brain derived antigen. A full validation of the protein using a large panel of serum samples from confirmed CCHF patients is now required. The assay was shown to be reproducible and stability studies showed that although the OD signal decreased with time, the low positive control was still detected 16 weeks after preparation of the plates. In addition, we were able to store the protein frozen at -70ºC indefinitely prior to denaturation and purification (data not shown).

In conclusion we were unable to express the CCHFV NP using the native gene of the S segment cloned in pQE-80L which employs a T5 promotor (results shown), or using the pCold TF expression system under the cold shock promoter (unpublished). Although there are no absolute rules that can be used to predict whether a protein will be expressed from a heterologous host it is possible to optimize a gene so that the codons are modified to take into account the codon bias of a particular host. Whether the codon optimization is responsible for the improved protein expression and yield that has been reported for various recombinant antigens or whether the optimization has altered secondary structure or another undetermined factor that influences the expression is not clear (Bradel-Tretheway et al., 2003; Li et al., 2008; Han et al., 2009; Simabuco et al., 2009). In addition to codon optimization, we selected to use Takara′s pCold TF DNA Vector which is a fusion cold shock expression vector that expresses Trigger Factor (TF) chaperone as a soluble tag.

Optimizing the codons for expression in *E. coli* improved both CAI and F_{OP} values generally associated with enhanced protein expression of in *E.coli*. The outcome was a protein that was readily expressed under standard conditions recommended by the suppliers of the pCOLD-TF vector when compared with no protein expression using the native gene. We were able to express a high yield of recombinant protein that induced a detectable immune response in mice and was biologically active in immmunoassays. Hence the results suggest that bacterially expressed proteins lacking post translational modifications and folding that occur with mammalian and baculovirus expression systems have potential for use in both diagnosis and surveillance.

REFERENCES

Ahmed, J., Bouloy, M., Ergonul, O., Fooks, A., Paweska, J., Chevalier, V., Drosten, C., Moormann, R., Tordo, N., Vatansever, Z., Calistri, P., Estrada-Pena, A., Mirazimi, A., Unger, H., Yin, H., Seitzer, U., 2009. International network for capacity building for the control of emerging viral vector-borne zoonotic diseases: ARBO-ZOONET. Euro. Surveill. 14, 19160-19163.

Ahmed, J., McFalls, J.M., Hoffman, C., Filone, C.M., Stewart, S.M., Paragas, J., Khodjaev, S., Shermukhamedova, D., Schmaljohn, C.S., Doms, R.W., Bertolotti-Ciarlet, A., 2005. Presence of broadly reactive and group-specific neutralizing epitopes on newly described isolates of Crimean-Congo hemorrhagic fever virus. J. Gen. Virol. 89, 3327-3336.

Al Tikriti, S.K., Al ani, F., Jurji, F.J., Tantawi, H., Al Moslih, M., Al Janabi, N., Mahmud, M.I., Al Bana, A., Habib, H., Al Munthri, H., Al Janabi, S., Al Jawahry, K., Yonan, M., Hassan, F., Simpson, D.I., 1981. Congo/Crimean haemorrhagic fever in Iraq. Bull. World Health Organ. 59, 85-90.

Baneyx, F., 1999. Recombinant protein expression in *Escherichia coli*. Curr. Opin. Biotechnol. 1999. 10, 411-421.

Baskerville, A., Satti, A., Murphy, F.A., Simpson, D.I., 1981. Congo-Crimean haemorrhagic fever in Dubai: histopathological studies. J. Clin. Pathol. 34, 871-874.

Bradel-Tretheway, B.G., Khen, Z., Dewhurst, S., 2003. Effects of codon-optimization on protein expression by the human herpesvirus 6 and 7 U51 open reading frame. J Virol. Methods 111, 145-156.

Briese, T., Paweska, J.T., McMullan, L.K., Hutchison, S.K., Street, C., Palacios, G., Khristova, M.L., Weyer, J., Swanepoel, R., Egholm, M., Nichol, S.T., Lipkin, W.I., 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever–associated arenavirus from southern Africa. Emerg. Infect. Dis. 5, 1-8.

Bui, H.H., Sidney, J., Li, W., Fusseder, N., Alessandro, S., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinformatics 8, 361-366.

Burney, M.I., Ghafoor, A., Saleen, M., Webb, P.A., Casals, J., 1980. Nosocomial outbreak of viral hemorrhagic fever caused by Crimean hemorrhagic fever –Congo virus in Pakistan, January 1976. Am. J. Trop. Med. Hyg. 29, 941-947.

Burt, F.J., 2011. Laboratory diagnosis of Crimean-Congo haemorrhagic fever. Future Virol. In press.

Burt F.J., Leman, P.A., Abbott, J.C., Swanepoel, R., 1994. Serodiagnosis of Crimean-Congo haemorrhagic fever. Epidemiol. Infect. 111, 547-557.

Burt, F.J., Leman, P.A., Smith, J.F., Swanepoel, R., 1998. The use of a reverse transcription-polymerase chain reaction for the detection of viral nucleic acid in the

diagnosis of Crimean-Congo hemorrhagic fever. J. Virol. Methods 70, 129-137.

Burt, F.J., Paweska, J.T., Ashkettle, B., Swanepoel, R., 2009. Genetic relationship among southern African Crimean-Congo haemorrhagic fever virus isolates and evidence for occurrence of reassortment. Epidemiol. Infect. 1379, 1302-1308.

Burt, F.J., Spencer, D.C., Leman, P.A., Patterson, B., Swanepoel, R., 1996. Investigation of tick-borne viruses as pathogens of humans in South Africa and evidence of Dugbe virus infection in a patient with prolonged thrombocytopenia. Epidemiol. Infect. 116, 353-361.

Burt, F.J., Swanepoel, R. 2005. Molecular epidemiology of African and Asian Crimean-Congo haemorrhagic fever isolates. Epidemiol. Infect. 4, 659-666.

Burt, F.J., Swanepoel, R., Braack, L.E.O., 1993. Enzyme-linked immunosorbent assays for the detection of antibody to Crimean-Congo haemorrhagic fever virus in the sera of livestock and wild vertebrates. Epidemiol. Infect. 111, 547-557.

Burt, F.J., Swanepoel, R., Shieh, W.J., Smith, J.F., Leman, P.A., Greer, P.W., Coffield, L.M., Rollin, P.E., Ksiazek, T.G., Peters, C.J., Zaki, S.R., 1997. Immunohistochemical and in situ localization of Crimean-Congo hemorrhagic fever (CCHF) virus in human tissues and implications for CCHF pathogenesis. Arch. Pathol. Lab. Med. 121, 839-846.

Butenko, A.M., Chumakov, M.P., 1990. Isolation of Crimean-Congo hemorrhagic fever virus from patients and from autopsy specimens. Arch. Virol. 1, 295-301.

Calisher, C.H., Karabatsos, N., 1989. Arbovirus serogroups: definition and geographic distribution, in: Monath, T.P. (Eds.), The Arboviruses: Epidemiology and Ecology. Vol 1. CRC Press Inc., Florida, pp.19-57.

Camicas, J.L., Wilson, M.L., Cornet, J.P., Digoutte, J.P., Calvo, M.A., Adam, F., Gonzalez, J.P., 1991. Ecology of ticks as potential vectors of Crimean-Congo hemorrhagic fever virus in Senegal: epidemiological implications. Arch. Virol. S1, 303.

Capua, I., 1998. Crimean-Congo haemorrhagic fever in ostriches: a public health risk for countries of the European Union. Avian Pathol. 27, 117-120.

Casals, J., 1969. Antigenic similarity between the virus causing Crimean hemorrhagic fever and Congo virus. Proc. Soc. Exp. Biol. Med. 131, 233-236.

Casals, J., Tignor, G.H., 1980. The *Nairovirus* genus: serological relationships. Intervirology 14, 144-147.

Chumakov, M.P., Smirnova, S.E., Tkachenko, E.A., 1970. Relationship between strains of Crimean haemorrhagic fever and Congo viruses. Acta Virol. 14, 82-85.

Clerx, J.P.M., Casals, J., Bishop, D.H.L., 1981. Structural characteristics of nairoviruses (genus *Nairovirus, Bunyaviridae*). J. Gen. Virol. 55, 165-178.

Davies, F.G., Casals, J., Jesset, D.M., Ochieng, P., 1978. The serological relationships of Nairobi sheep disease virus. J. Comp. Pathol. 88, 519-523.

Deyde, V.M., Khristova, M. L., Rollin, P.E., Ksiazek, T.G., Nichol, S.T., 2006. Crimean-Congo hemorrhagic fever virus genomics and global diversity. J. Virol. 80, 8834-8842.

Donets, M.A., Rezapkin, G.V., Ivanov, A.P. Tkachenko, E.A., 1982. Immunosorbent assays for the diagnosis of Crimean-Congo hemorrhagic fever (CCHF). Am. J. Trop. Med. Hyg. 31,156-162.

Drosten, C., Gottig, S., Schilling, S., Asper, M., Panning, M., Schmitz, H., Gunther, S., 2002. Rapid detection and quantification of RNA of Ebola and Marburg viruses, Lassa virus, Crimean–Congo hemorrhagic fever virus, Rift Valley fever virus, dengue virus, and yellow fever virus by real-time reverse transcription-PCR. J. Clin. Microbiol. 40, 2323–2330.

Drosten, C., Kummerer, B.M., Schmitz, H., Gunther, S., 2003. Molecular diagnostics of viral hemorrhagic fevers. Antiviral Res. 57, 61–87.

Dunster, L., Dunster, M., Ofula, V., Beti, D., Kazooba-Voskamp, F., Burt, F., Swanepoel, R., DeCock, K.M., 2002. First documentation of human Crimean-Congo hemorrhagic fever, Kenya. Emerg. Infect. Dis. 8, 1005-1006.

Durden, L.A., Logan, T.M., Wilson, M.L., Linthicum, K.J., 1993. Experimental vector incompetence of a soft tick, Ornithodoros sonrai (Acari: Argasidae), for Crimean-Congo hemorrhagic fever virus. J. Med. Entomol. 30, 493-496.

El Azazy, O.M., Scrimgeour, E.M., 1997. Crimean-Congo haemorrhagic fever virus infection in the western province of Saudi Arabia. Trans. R. Soc. Trop. Med. Hyg. 91, 275-278.

Elliot, R.M., 1990. Molecular biology of the *Bunyaviridae*. J. Gen. Virol. 71, 501-522.

Ergonul, O., 2006. Crimean-Congo haemorrhagic fever. Lancet Infect. Dis. 6, 203-214.

Ergonul, O., 2008. Treatment of Crimean-Congo hemorrhagic fever. Antiviral Res. 78, 125-131.

Fink, A.L., 1998. Protein aggregation: folding aggregates, inclusion bodies and amyloid. Fold. Des.. 3, 9-23.

Fisher-Hoch, S.P., Khan, J.A., Rehman, S., Mirza, S., Khurshid, M., McCormick, J.B., 1995. Crimean Congo-haemorrhagic fever treated with oral ribavirin. Lancet 346, 472–475.

Frey, A., Di Canzio, J., Zurakowski, D., 1998. A statistically defined endpoint titer determination method for immunoassays. J. Immunol. Methods 221, 35-41.

Garcia, S., Chinikar, S., Coudrier, D., Billecocq, A., Hooshmand, B., Crance, J.M., Garin, D., Bouloy, M., 2006. Evaluation of a Crimean-Congo hemorrhagic fever recombinant antigen expressed by Semliki Forest suicide virus for IgM and IgG antibody detection in human and animal sera collected in Iran. J. Clin. Microbiol. 35, 154-159.

Gear, J.H., Thompson, P.D., Hopp, M., Andronikou, S., Cohn, R.J., Ledger, J., Berkowitz, F.E., 1982. Congo-Crimean haemorrhagic fever in South Africa: report of a fatal case in the Transvaal. S. Afr. Med. J. 62, 576-580.

Gerdes, G.H., 2004. Rift Valley fever. Rev. Sci.Tech. 23, 613-623.

Graci, J.D., Cameron, C.E., 2006. Mechanisms of action of ribavirin against distinct viruses. Rev. Med. Virol. 16, 37–48.

Gustafsson, C., Govindarajan, S., Minshull, J., 2004. Codon bias and heterologous protein expression. Trends Biotechnol. 22, 346-353.

Hayden, F.G., 2006. Antiviral agents, in: Brunton, L.L. (Eds.), Goodman and Gilman's: The Pharmacological Basis of Therapeutics. McGraw-Hill, pp. 1265–1267.

Han, J.H., Choi, Y.S., Kim, W.J., Jeon, Y.H., Lee, S.K., Lee, B.J., Ryu, K.S., 2009. Codon optimization enhances protein expression of human peptide deformylase in *E.coli*. Protein Expr. Purif. 70, 224-230.

Henry, I., Sharp,P.M., 2007. Predicting Gene Expression Level from Codon Usage Bias. Mol. Biol. Evol. 24, 10-12.

Hewson, R., Chamberlain, J., Mioulet, V., Lloyd, G., Jamil, B., Hasan, R., Gmyl, A., Gyml, L., Smirnova, S.E., Lukashev, A., Karganova, G., Clegg, C., 2004a. Crimean-Congo haemorrhagic fever virus: sequence analysis of the small RNA segments from a collection of viruses worldwide. Virus Res. 102, 185-189.

Hewson, R., Gmyl, A., Gmyl, L., Smirnova, S.E., Karganova, G., Jamil, B., Hassan, R., Chamberlain, J., Clegg, C., 2004b. Evidence of segment reassortment in Crimean-Congo haemorrhagic fever virus. J. Gen. Virol. 85, 3059-3070.

Hoogstraal, H., 1979. The epidemiology of tick-borne Crimean-Congo hemorrhagic fever in Asia, Europe and Africa. J. Med. Entomol. 15, 307-417.

Hoogstraal, H., Kaiser, M.N., Traylor, M.A., Gaber, S., Guindy, E., 1961. Ticks (Ixodoidea) on birds migrating from Africa to Europe and Asia. Bull. World Health Organ. 24, 197-212.

Jacobson, R.H., 1998. Validation of serological assays for diagnosis of infectious diseases. Rev. Sci. Tech. 17, 469-526.

Jana, S., Deb, J.K., 2005. Strategies for efficient production of heterologous proteins in *Escherichia coli*. Appl. Microbiol. Biotechnol. 67, 289-298.

Joubert, J.R., King, J.B., Rossouw, D.J., Cooper, R., 1985. A nosocomial outbreak of Crimean-Congo haemorrhagic fever at Tygerberg Hospital. Part III. Clinical pathology and pathogenesis. S. Afr. Med. J. 68, 722-728.

Kane, J.F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr. Opin. Biotechnol. 6, 494-500.

Karabatsos, N., 1985. International Catalogue of Arboviruses (including certain other viruses of vertebrates), third ed. Am. Soc. Trop. Med. Hyg., San Antonio.

Karti, S.S., Odabasi, Z., Korten, V., Yilmaz, M., Sonmez, M., Caylan, R., Akdogan, E., Eren, N., Koksal, I., Ovali, E., Erickson, B.R., Vincent, M.J., Nichol, S.T., Comer, J.A., Rollin, P.E., Ksiazek, T.G., 2004. Crimean-Congo hemorrhagic fever in Turkey. Emerg. Infect. Dis. 10, 1379-1384.

Kondiah, K., Swanepoel, R., Paweska, J.T., Burt, F.J., 2010. A Simple-Probe real-time PCR assay for genotyping reassorted and non-reassorted isolates of Crimean-Congo hemorrhagic fever virus in southern Africa. J. Virol. Methods 169, 34-38.

Kondratenko, V.F., 1976. Importance of ixodid ticks in the transmission and preservation of the causative agent of Crimean hemorrhagic fever in the foci of the infection. Parazitologiia 10, 297-302.

Kuhn, J.H., Seregin, S.V., Morzunov, S.P., Petrova, I.D., Vyshemirskii, O.I., Lvov, D.K., Tyunnikov, G.I., Gutorov, V.V., Netesov, S.V., Petrov, V.S., 2004. Genetic analysis of the mRNA segment of Crimean-Congo haemorrhagic fever virus involved in the recent outbreaks in Russia. Arch. Virol. 14911, 2199-2213.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Li, Z., Hong, G., Wu, Z., Hu, B., Xu, J., Li, L., 2008. Optimization of the expression of hepatits B virus gene in Pichia pastoris and immunological characterization of the product. J. Biotechnol. 138, 1-8.

Lilie, H., Schwarz, E., Rudolph, R., 1998. Advances in refolding of proteins produced in *E. coil.* Curr. Opin. Biotechnol. 9, 497-501.

Lithwick, G., Margalit, H., 2003. Hierarchy of sequence-dependent features associated with prokaryotic translation. Genome Res. 13, 2665-2673.

Luckow, V. A. 1991. Cloning and expression of heterologous genes in insect cells with baculovirus vectors, in: Prokop, A., Bajpai, R.K. (Eds.), Recombinant DNA technology and applications. McGraw-Hill., New York, pp. 97–152.

Magurano, F., Nicoletti, L., 1999. Humoral response in Toscana virus acute neurological disease investigated by viral-protein-specific immunoassays. Clin. Diag. Lab. Immunol. 6, 55-60.

Maltezou, H.C., Andonova, L., Andraghetti, R., Bouloy, M., Ergonul, O., Jongejan, F., Kalvatchev, N., Nichol, S., Niedrig, M., Platonov, A., Thomson, G., Leitmeyer, K., Zeller., 2010. Crimean-Congo hemorrhagic fever in Europe: current situation calls for preparedness. Euro. Surveill. 15, 19504-19508.

Maltezou, H.C., Papa, A., 2010. Crimean-Congo hemorrhagic fever: risk for emergence of new endemic foci in Europe? Travel Med. Infect. Dis. 8, 139-143.

Marriott, A.C., Nuttall, P.A., 1992. Comparison of the S RNA segments and nucleoprotein sequences of Crimean-Congo hemorrhagic fever, Hazara, and Dugbe viruses. Virology. 189, 795-799.

Marriott, A.C., Nuttall, P.A., 1996. Molecular biology of nairoviruses, in: Elliott, R.M., (Eds.), The *Bunyaviridae*. Plenum Press., New York.

Marriott, A.C., Polyzoni, T., Antoniadis, A., Nuttall, P.A., 1994. Detection of human antibodies to Crimean-Congo haemorrhagic fever virus using expressed viral nucleocapsid protein. J. Gen. Virol. 75, 2157-2161.

McCormick, J.B., King, I.J., Webb, P.A., Scribner, C.L., Craven, R.B., Johnson, K.M., Elliott, L.H., Belmont-Williams, R., 1986. Lassa fever: effective therapy with ribavirin. N. Engl. J. Med. 314, 20–26.

Middelberg, A. P. J., 2004. Preparative protein folding. Trends Biotechnol. 20, 433–437.

Miller, L.K., 1988. Baculoviruses as gene expression vectors. Annu. Rev. Microbiol. 42, 177–199.

Morikawa, S., Saijo, M., Kurane, I., 2007. Recent progress in molecular biology of Crimean-Congo hemorrhagic fever. Comp. Immunol. Microbiol. Infect. Dis. 30, 375-389.

Papa, A., Dalla, V., Papadimitriou, E., Kartalis, G.N., Antoniadis, A., 2010. Emergence of Crimean-Congo haemorrhagic fever in Greece. Clin. Microbiol. Infect. 16, 843-847.

Papa, A., Ma, B., Kouidou, S., Tang, Q., Hang, C., Antoniadis, A., 2002. Genetic characterization of the M RNA segment of Crimean-Congo hemorrhagic fever virus strains, China. Emerg. Infect. Dis. 8, 50-53.

Paweska, J.T., Sewlall, N.H., Ksiazek, T.G., Blumberg, L.H., Hale, M.J., Lipkin, W.I., Weyer, J., Nichol, S.T., Rollin, P.E., McMullan, L.K., Paddock, C.D., Briese, T., Mnyaluza, J., Dinh, T.H., Mukonka, V., Ching, P., Duse, A., Richards, G., de Jong, G., Cohen, C., Ikalafeng, B., Mugero, C., Asomugha, C., Malotle, M.M., Nteo, D.M., Misiani, E., Swanepoel, R., Zaki, S.R., 2009. Nosocomial outbreak of novel arenavirus infection, southern Africa. Emerg. Infect. Dis. 15, 1598-1602.

Puigbo, P., Guzman, E., Romeu, A., Garcia-Vallvé, S., 2007. Optimizer: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res. 10, 1-6.

Randolph, S.E., Rigers, D.J., 2007. Ecology of tick-borne disease and the role of climate, in: Ergonul, O., Whitehouse, C.A. (Eds.), Crimean-Congo haemorrhagic fever: a global perspective. Dordrecht Springer., pp. 1167-1186.

Rocha, P.C., 2004. Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. Genome Res. 14, 2279-2286.

Rodriguez, L.L., Maupin, G.O., Ksiazek, T.G., Rollin, P.E., Khan, A.S., Schwarz, T.F., Lofts, R.S., Smith, J.F., Noor, A.M., Peters, C.J., Nichol, S., 1997. Molecular investigation of a multisource outbreak of Crimean-Congo hemorrhagic fever in the United Arab Emirates. Am. J. Trop. Med. Hyg. 57, 512-518.

Sahdev, S., Khattar, S.K., Saini, K.S., 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol. Cell. Biochem. 307, 249–264.

Saijo, M., Tang, Q., Niikura, M., Maeda, A., Ikegami, T., Prehaud, C., Kurane, I., Morikawa, S., 2002a. Recombinant nucleoprotein-based enzyme-linked immunosorbent assay for detection of immunoglobulin G antibodies to Crimean-Congo Hemorrhagic Fever Virus. J. Clin. Microbiol. 40, 1587-1591.

Saijo, M., Tang, Q., Niikura, M., Maeda A., Ikegami, T., Sakai, K., Prehaud, C., Kurane, I., Morikawa, S., 2002b. Immunofluorescence technique using HeLa cells expressing recombinant nucleoprotein for detection of immunoglobulin G antibodies to Crimean-Congo hemorrhagic fever virus. J. Clin. Microbiol. 40, 372-375.

Saijo, M., Tang., Q, Shimayi, B., Han, L., Zhang, Y., Asiguma, M.,Tianshu, D.,Maeda, A., Kurane, I., Morikawa, S., 2005. Recombinant nucleoprotein-based serological diagnosis of Crimean-Congo hemorrhagic fever. J. Med. Virol. 75, 295-299.

Saluzzo, J.F., Aubry, F.P., McCormick, J., Digoutte, J.P., 1985. Haemorrhagic fever caused by Crimean-Congo haemorrhagic fever virus in Mauritania. Trans. R. Soc. Trop. Med. Hyg. 79, 268.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: a laboratory manual, thirded. Cold Spring Harbor, New York.

Sanchez, A.J., Vincent, M.J., Nichol, S.T., 2002. Characterization of the glycoproteins of Crimean-Congo hemorrhagic fever virus. J. Virol. 76, 7263-75.

Schmaljohn, C.S., Patterson, J.L., 1990. Bunyaviridae and their replication. Part II: Replication of *Bunyaviridae*, in: Fields, B.N., Knipe, D.M., Howley, P.M. (Eds.), Virology, second ed. Raven Press Ltd., New York, pp. 1147-1173.

Schwarz, T.F., Nitschko, H., Jager, G., Nsanze, H., Longson, M., Pugh, RN., Abraham, A.K., 1995. Crimean-Congo haemorrhagic fever in Oman. Lancet 346, 1230.

Schwarz, T.F., Nsanze, H., Longson, M., H., Gilch, S., Shurie, H., Ameen, A., Zahir,

A.R., Acharya, U.G., Jager, G., 1996. Polymerase chain reaction for diagnosis and identification of distinct variants of Crimean-Congo hemorrhagic fever virus in the United Arab Emirates. Am. J. Trop. Med. Hyg. 55, 190-6.

Sharp, P.M., Bailes, E., Grocock, R.J., Peden, J.F., Sockett, R.E., 2005. Variation in the strength of selected codon bias usage among bacteria. Nucleic Acids Res. 33, 1141-1153.

Sharp, P.M., Li, W.H., 1987. The codon adaptation index – a measure of directional synonomous codon usage bias, and its potential applications. Nucleic Acids Res. 15, 1281-1295.

Shepherd, A.J., Leman, P.A., Swanepoel, R., 1989a. Viremia and antibody response of small African and laboratory animals to Crimean-Congo hemorrhagic fever virus infection. Am. J. Trop. Med. Hyg. 40, 541-547.

Shepherd, A.J., Swanepoel, R., Gill., DE., 1988. Evaluation of enzyme-linked immunosorbent assay and reversed passive hemagglutination for detection of Crimean-Congo hemorrhagic fever virus. J. Clin. Microbiol. 26, 347-353.

Shepherd, A.J., Swanepoel, R., Leman, P.A., 1989b. Antibody response in Crimean-Congo hemorrhagic fever. Rev. Infect. Dis. 11, 801-806.

Shepherd, A.J., Swanepoel, R., Leman, P.A., Shepherd, S.P., 1986. Comparison of methods for isolation and titration of Crimean-Congo hemorrhagic fever virus. J. Clin. Microbiol. 24, 654-656.

Shepherd, A.J., Swanepoel, R., Shepherd, S.P., Leman, P.A., Mathee, O., 1991. Viraemic transmission of Crimean-Congo haemorrhagic fever virus to ticks. Epidemiol. Infect. 106, 373-382.

Shepherd, A.J., Swanepoel, R., Shepherd, S.P., McGillivray, G.M., Searle, L.A., 1987. Antibody to Crimean-Congo hemorrhagic fever virus in wild mammals from southern Africa. Am. J. Trop. Med. Hyg. 36, 133-142.

Simabuco, F.M., Tamura, R.E., Carromeu, C., Farinha-Arcieri, L.E., Ventura, A.M., 2009. Gene optimization leads to rocust expression of human respiratory syncytial virus nucleoprotein and phosphoprotein in human cells and induction of humoral immunity in mice. J. Virol. Methods 158, 93-99.

Simmons, J.H., 2008. Development, application, and quality control of serology assays used for diagnostic monitoring of laboratory nonhuman primates. ILAR J. 49, 157-169.

Singh, S.M., Panda, A.K., 2005. Solubilization and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. 99, 303-310.

Sorensen, H.P., Mortensen, K.K., 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J. Biotechnol. 115, 113-128.

Suleiman, M.N., Muscat-Baron, J.M., Harries, J.R., Satti, A.G., Platt, G.S., Bowen, E.T., Simpsom, D.I., 1980. Congo/Crimean haemorrhagic fever in Dubai. An outbreak at the Rashid Hospital. Lancet 2, 939-941.

Swanepoel, R., Gill, D.E., Shepherd, A.J., Leman, P.A., Mynhardt, J.H., Harvey, S., 1989. The clinical pathology of Crimean-Congo hemorrhagic fever. Rev. Infect. Dis. 11, 794-800.

Swanepoel, R., Leman, P.A., Burt, F.J., Jardine, J., Verwoerd, D.J., Capua, I., Bruckner, G.K., Burger, W.P., 1998. Experimental infection of ostriches with Crimean-Congo haemorrhagic fever virus. Epidemiol. Infect. 121, 427-432.

Swanepoel, R., Shepherd, A.J., Leman, P.A., Shepherd, S.P., McGillivray, G.M., Erasmus, M.J., Searle, L. A., Gill, D.E., 1987. Epidemiologic and clinical features of Crimean-Congo hemorrhagic fever in southern Africa. Am. J. Trop. Med. Hyg. 36, 120-132.

Swanepoel, R., Struthers, J.K., McGillivray, G.M., 1983. Reversed passive hemagglutination and inhibition with Rift Valley fever and Crimean-Congo hemorrhagic fever viruses. Am. J. Trop. Med. Hyg. 32, 610-617.

Tang, Q., Saijo, M., Zhang, Y., Asiguma, M., Tianshu, D., Han, L., Shimayi, B., Maeda, A., Kurane, I., Morikawa, S., 2003. A patient with Crimean-Congo hemorrhagic fever serologically diagnosed by recombinant nucleoprotein-based antibody detection systems. Clin. Diagn. Lab. Immunol. 10, 489-491.

Tignor, G.H., Hanham, C.A., 1993. Ribavirin efficacy in an in vivo model of Crimean-Congo hemorrhagic fever virus (CCHF) infection. Antiviral Res. 22, 309–325.

Tsumoto, K., Ejima, D., Kumagai, I., Arakawa, T., 2003. Practical considerations in refolding proteins from inclusion bodies. Protein Expr. Purif. 28, 1-8.

Vorou, R., Pierroutsakos, I.N., Maltezou, H.C., 2007. Crimean-Congo hemorrhagic fever. Curr. Opin. Infect. Dis. 20, 495-500.

Ward, K.W., Marriott, A.C., El-Ghorr, A.A., Nuttall., 1990. Coding strategy of the SRNA segment of Dugbe virus (Nairovirus, Bunyaviridae). Virology 175, 518-524.

Watts, D.M., Ksiazek, T.G., Linthicum, K.J., Hoogstraal, H., 1989a. Crimean-Congo haemorrhagic fever, in: Monath, T.P., (Eds.), The Arboviruses: Epidemiology and Ecology. Vol II. CRC Press Inc., Florida, pp. 177-222.

Watts, D.M., Ussery, M.A., Nash, D., Peters, C.J., 1989b. Inhibition of Crimean-Congo hemorrhagic fever viral infectivity yields in vitro by ribavirin. Am. J. Trop. Med. Hyg. 41, 581–585.

Whitehouse, A., 2004. Crimean-Congo hemorrhagic fever. Antiviral Res. 64, 145-160.

Wright, F., 1990. The ‘effective number of codons’ used in a gene. Gene 87, 23-29.

Zavitsanou, A., Babatsikou, F., Koutis, C., 2009. Crimean-Congo hemorrhagic fever: an emerging tick-borne disease. Health Sci. J. 3, 10-18.

INTERNET REFERENCES

Center for Disease Control (CDC) . Internet. Available from http://www.dpd.cdc.gov/dpdx

World Health Organization (WHO). Internet. Available from http://www.dpd.cdc.gov/dpdx

APPENDIX 1

ELISA RAW DATA

Table 17. Raw data of checkerboard titration using two- fold dilutions of recombinant CCHFV NP

	1:500	1:500	1:1000	1:1000	1:2000	1:2000	1:4000	1:4000	
CCHFV NP	2.469	2.36	1.927	2.046	1.463	1.353	1.046	0.8	**C++**
CCHFV NP	0.154	0.102	0.088	0.076	0.077	0.07	0.059	0.067	**C-**
MOCK	0.175	0.129	0.126	0.201	0.165	0.13	0.106	0.13	**C++**
MOCK	0.111	0.07	0.07	0.068	0.053	0.065	0.073	0.052	**C-**

Table 18. Raw data of IgG ELISA for repeatability study assessed over 4 days using recombinant CCHFV NP diluted 1:2000

	70/08		**67/08**		**57/09**		**10/09**		**C-**		
CCHFV NP	0.924	1.077	0.934	0.655	0.507	0.68	0.579	0.563	0.2560	0.2480	**Day 1**
MOCK	0.089	0.092	0.142	0.105	0.048	0.058	0.091	0.086	0.0740	0.0760	
	70/08		**67/08**		**57/09**		**10/09**		**C-**		
CCHFV NP	0.755	0.687	0.981	0.808	0.469	0.525	0.665	0.532	0.302	0.181	**Day 2**
MOCK	0.113	0.098	0.151	0.101	0.051	0.042	0.077	0.066	0.121	0.079	
	70/08		**67/08**		**57/09**		**10/09**		**C-**		
CCHFV NP	0.837	0.948	0.868	0.773	0.508	0.441	0.53	0.439	0.318	0.2723	**Day 3**
MOCK	0.093	0.106	0.142	0.103	0.045	0.042	0.075	0.066	0.172	0.106	
	70/08		**67/08**		**57/09**		**10/09**		**C-**		
CCHFV NP	0.967	0.944	1.023	0.897	0.471	0.475	0.588	0.573	0.318	0.223	**Day 4**
MOCK	0.12	0.122	0.154	0.165	0.119	0.091	0.105	0.093	0.172	0.106	

Table 19. Raw data of IgG ELISA for detection of antibodies to CCHFV in convalescent patients, using recombinant CCHFV NP diluted 1:2000

	6/08		71/08		53/10		10/08		11/08		C-	
CCHFV	1.4120	1.5930	1.455	1.294	1.2430	1.2720	0.49	0.412	0.447	0.434	0.367	0.382
MOCK	0.2010	0.2120	0.242	0.201	0.1170	0.11	0.11	0.104	0.126	0.126	0.146	0.16
	70/08		52/10		22/08		66/08		42/10		67/08	
CCHFV	0.919	1.064	1.2720	1.294	0.401	0.495	1.496	1.264	1.138	1.048	0.834	0.842
MOCK	0.108	0.092	0.2050	0.102	0.058	0.04	0.132	0.102	0.114	0.093	0.089	0.08
	41/10		69/08		6/09		10/09		57/09		C-	
CCHFV	0.725	0.735	0.965	0.893	1.082	1.031	0.602	0.733	0.476	0.428	0.146	0.223
MOCK	0.074	0.075	0.106	0.107	0.09	0.067	0.074	0.073	0.067	0.067	0.16	0.151
	53/09		30/10		29/10		51/10		6/11		29/11	
CCHFV	1.193	1.154	1.432	1.736	1.5850	1.264	0.434	0.441	1.046	0.9	0.969	1.057
MOCK	0.129	0.118	0.108	0.107	0.1560	0.102	0.08	0.086	0.14	0.126	0.065	0.075

Table 20. Raw data of IgG ELISA for assessment of stability over 4 months, using recombinant CCHFV NP diluted 1:2000

		O/N		Week 2		Week 4		Week 12		Week 16	
CCHFV	**C++**	1.898	1.616	1.862	1.81	0.961	0.892	0.61	0.565	0.385	0.727
	C+	0.674	0.652	0.727	0.556	0.57	0.543	0.306	0.343	0.481	0.437
	C-	0.303	0.3185	0.223	0.1925	0.162	0.223	0.167	0.196	0.329	0.363
MOCK	**C++**	0.184	0.172	0.284	0.247	0.076	0.086	0.094	0.103	0.113	0.104
	C+	0.087	0.069	0.104	0.1085	0.112	0.1825	0.068	0.075	0.08	0.086
	C-	0.074	0.0755	0.074	0.0755	0.077	0.074	0.077	0.085	0.236	0.216

Table 21. Raw data of IgG ELISA performed on negative panel to determine cut off, using recombinant CCHFV NP diluted 1:2000

	17/07	20/07	21/07	22/07	23/07	25/07	18/08	C++	C++	
CCHFV NP	0.24	0.275	0.429	0.093	0.319	0.182	0.164	1.229	1.044	**Day 1**
MOCK	0.16	0.093	0.372	0.083	0.134	0.084	0.12	0.193	0.149	
	19/08	**28/08**	**43/09**	**44/09**	**47/09**	**48/09**	**50/09**	**C++**	**C++**	
CCHFV NP	0.248	0.124	0.197	0.339	0.302	0.181	0.223	1.085	1.004	**Day 1**
MOCK	0.111	0.081	0.108	0.099	0.121	0.079	0.106	0.226	0.226	
	17/07	**20/07**	**21/07**	**22/07**	**23/07**	**25/07**	**18/08**	**C++**	**C++**	
CCHFV NP	0.154	0.162	0.445	0.07	0.231	0.152	0.12	0.977	0.979	**Day 2**
MOCK	0.09	0.07	0.57	0.044	0.079	0.055	0.042	0.056	0.051	
	19/08	**28/08**	**43/09**	**44/09**	**47/09**	**48/09**	**50/09**	**C++**	**C++**	
CCHFV NP	0.189	0.198	0.256	0.248	0.223	0.13	0.16	0.969	1.059	**Day 2**
MOCK	0.083	0.055	0.074	0.076	0.112	0.045	0.059	0.056	0.062	
	17/07	**20/07**	**21/07**	**22/07**	**23/07**	**25/07**	**18/08**	**C++**	**C++**	
CCHFV NP	0.244	0.165	0.438	0.067	0.283	0.252	0.193	1.046	1.298	**Day 3**
MOCK	0.106	0.105	0.49	0.041	0.085	0.059	0.051	0.051	0.072	
	19/08	**28/08**	**43/09**	**44/09**	**47/09**	**48/09**	**50/09**	**C++**	**C++**	
CCHFV NP	0.214	0.204	0.319	0.246	0.248	0.151	0.298	1.058	1.261	**Day 3**
MOCK	0.084	0.05	0.096	0.08	0.177	0.06	0.061	0.054	0.067	
	17/07	**20/07**	**21/07**	**22/07**	**23/07**	**25/07**	**18/08**	**C++**	**C++**	
CCHFV NP	0.189	0.166	0.445	0.079	0.175	0.141	0.12	1.346	1.252	**Day 4**
MOCK	0.113	0.074	0.529	0.05	0.085	0.067	0.065	0.077	0.11	

	19/08	28/08	43/09	44/09	47/09	48/09	50/09	C++	C++	
CCHFV NP	0.213	0.174	0.138	0.145	0.242	0.138	0.172	1.243	1.214	**Day 4**
MOCK	0.174	0.099	0.105	0.099	0.195	0.071	0.103	0.088	0.1	

Table 22. Raw data of IgG ELISA performed to determine antibody titers of convalescent sera using recombinant CCHFV NP diluted 1:2000, plate 1

	Dilution	6/08	66/08	67/08	69/08	70/08	71/08	6/09	10/09	57/09	29/10	C-	C-
CCHFV	**1:100**	1.412	0.447	1.16	1.394	1.451	1.653	1.298	0.44	0.338	1.585	0.316	0.286
MOCK	**1:100**	0.201	0.108	0.161	0.166	0.11	0.202	0.122	0.102	0.153	0.156	0.139	0.156
CCHFV	**1:400**	0.729	0.208	0.439	0.704	0.861	0.991	0.898	0.17	0.135	0.85	0.102	0.113
MOCK	**1:400**	0.119	0.049	0.075	0.064	0.056	0.12	0.04	0.044	0.077	0.075	0.077	0.069
CCHFV	**1:1600**	0.329	0.097	0.122	0.181	0.288	0.441	0.325	0.062	0.049	0.404	0.045	0.053
MOCK	**1:1600**	0.069	0.04	0.043	0.05	0.042	0.061	0.03	0.038	0.051	0.045	0.041	0.051
CCHFV	**1:6400**	0.12	0.039	0.05	0.068	0.103	0.135	0.114	0.035	0.031	0.127	0.032	0.033
MOCK	**1:6400**	0.064	0.029	0.032	0.03	0.028	0.042	0.026	0.03	0.027	0.028	0.039	0.032

Table 23. Raw data of IgG ELISA performed to determine antibody titers of convalescent sera using recombinant CCHFV NP, plate 2

	Dilution	30/10	41/10	42/10	51/10	52/10	53/10	6/11	10/08	11/08	22/08	C-	C-
CCHFV	**1:100**	1.193	1.191	0.736	0.489	1.272	1.243	0.429	0.49	0.447	0.401	0.265	0.271
MOCK	**1:100**	0.129	0.205	0.136	0.112	0.205	0.117	0.086	0.11	0.126	0.058	0.14	0.17
CCHFV	**1:400**	0.875	0.495	0.32	0.18	0.597	0.748	0.328	0.12	0.127	0.136	0.114	0.111
MOCK	**1:400**	0.226	0.08	0.086	0.062	0.137	0.074	0.052	0.064	0.028	0.086	0.069	0.087
CCHFV	**1:1600**	0.364	0.174	0.084	0.064	0.268	0.416	0.109	0.051	0.048	0.045	0.046	0.05
MOCK	**1:1600**	0.098	0.048	0.046	0.038	0.072	0.045	0.055	0.035	0.034	0.033	0.046	0.053
CCHFV	**1:6400**	0.131	0.071	0.046	0.032	0.078	0.122	0.052	0.035	0.038	0.046	0.029	0.029
MOCK	**1:6400**	0.051	0.034	0.037	0.03	0.048	0.033	0.052	0.03	0.03	0.037	0.031	0.031

APPENDIX 2

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

ABTS = 2.2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid)
amp = ampicillin
C++ = high positive serum control
C+ = low positive serum control
C - = negative serum control
CCHFV = Crimean Congo Haemorrhagic Fever virus
cDNA = complimentary DNA
DNA = Deoxy-Ribonucleic acid
ddH_2O = double distilled water
ELISA = Enzyme Linked Immunosorbent Assay
HRPO = Horseradish peroxidise
IF = immunofluorescence
IFA – immunofluorescence assay
kDa = kilo Dalton
LB = Luria Bertani
NP = nucleoprotein
Ni^{2+}= nickel
NICD = National Institute for Communicable Disease
OD = Optical density
ORF = Open reading frame
PBS = Phosphate buffered saline
PCR = Polymerase Chain Reaction
PP = percentage positive
RNA = Ribonucleic acid
RT-PCR = Reverse Transcriptase PCR
SDS = sodium dodecyl sulphate
SDS-PAGE = SDS-Polyacrylamide gel electrophoresis
SPU = Special Pathogens Unit
TAE = Tris Acetate EDTA
TBS = Tris buffered saline
TE = tris-EDTA
TF = Trigger factor

APPENDIX 3

Title and abstract of paper to be submitted to Journal of Virological Methods for review

Gene optimization for bacterial expression of Crimean-Congo haemorrhagic fever virus nucleoprotein and induction of humoral immunity in mice

Rudo R Samudz[a]i, Robert Swanepoe[b]l, Janusz T Paweska[b], Patricia A Leman[b], Felicity Burt[a*]

Department of Medical Microbiology and Virology, National Health Laboratory Services Universitas and University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa.

Special Pathogens Unit, National Institute for Communicable Diseases, P Bag X 4 Sandringham, Johannesburg 2133, South Africa

Abstract

Crimean-Congo haemorrhagic fever virus (CCHFV) is a tick-borne viral zoonosis widely distributed in Africa, Asia, Russia and the Balkans. The emergence and re-emergence of CCHFV emphasizes the importance of increasing both human and veterinary surveillance and developing diagnostic capacity. Recombinant CCHFV nucleocapsid protein (NP) has been expressed using insect cells and mammalian cells and used as a diagnostic tool but bacterial expression has not yet previously been described. The S gene of CCHFV was codon optimized and the NP expressed in Escherichia coli from the synthetic gene. The protein was reacted against serum samples collected from confirmed patients at varying intervals after onset of illness from acute to convalescence stages using both an ELISA and Western blot. To confirm that the protein was able to induce a humoral antibody response that could be detected using CCHFV antigen derived from live virus, we immunized mice and tested serum samples using IF slides prepared from CCHFV infected Vero cells. The recombinant antigen was able to detect IgG antibody in acute and convalescent sera. In addition, we were able to induce a detectable IgG antibody response in mice immunized using the NP. The results suggest that bacterially expressed proteins lacking post translational modifications and folding that occur with mammalian and baculovirus expression can be used in ELISA to detect IgG antibody against CCHFV in human sera which finds application in diagnostics and epidemiologic and surveillance studies

*Corresponding author at: Department of Medical Microbiology and Virology, National Health Laboratory Services Universitas and University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa

Tel: 27 51 4053173
Email address: burtfj@ufs.ac.za (F.J. Burt)

APPENDIX 4

Title and abstract of presentation at the South African Society for Biochemistry and Molecular Biology (SASBMB), 18-20th January 2010, Bloemfontein

Preparation of recombinant antigens for demonstrating antibody responses in patients with Crimean-Congo haemorrhagic fever virus infections

Rudo Samudzi[a], Felicity Burt[b]

[a,b] Department of Medical Microbiology and Virology, National Health Laboratory Services Universitas, Faculty of Health Sciences, University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa

E-mail: rudosamudzi@yahoo.com

Keywords: CCHF, nucleoprotein, recombinant antigen
Topic: Biochemistry and Molecular Biology

Introduction: Crimean-Congo hemorrhagic fever (CCHF) virus causes a severe and often fatal infection. The virus has the propensity to cause nosocomial infections and hence a rapid and sensitive diagnosis is important for isolation of the patient for protection of health care workers and implementation of supportive therapy. Current serological diagnostic assays are based on enzyme linked immunosorbent assays (ELISA) or immunofluorescence (IF) tests using inactivated virus which necessitates biosafety level (BSL) 4 facilities for preparation of reagents. The aim of the study was to prepare a safe recombinant nucleoprotein and determine its suitability as a diagnostic tool for detection of antibody responses in survivors of CCHF infection.

Methods and materials: The DNA sequence encoding the nucleoprotein of CCHF was sent for codon optimization to facilitate protein expression from a bacterial expression system. The synthesized gene was cloned into a cold shock expression vector, pColdTF DNA vector and protein was expressed with an Ni^{+} terminal His Tag and Trigger Factor chaperone as a soluble tag. The protein was then purified using Ni^{+} columns. The recombinant antigen was used to develop an ELISA for detection of IgG antibodies against CCHF virus. A panel of known positive and negative serum samples will be used establish suitable cut off criteria and demonstrate repeatability of the assay.

Results: A 54kDa protein was co-expressed with Trigger Factor (52kDa). The purified protein was used to develop an ELISA which detected an IgG antibody response in convalescent serum samples from a confirmed CCHF patient. The ELISA will be validated using a panel of serum samples from confirmed patients to determine the specificity and sensitivity of the ELISA for diagnostic purposes.

Conclusion: Recombinant proteins have been shown to be safe, cost effective reagents that can be prepared for biohazardous pathogens without the requirements of BSL 4 facilities.

APPENDIX 5

Title and abstract of poster presented at the 13th International Congress on Infectious Diseases (ICID) 9 – 12th March 2010, Miami, USA

Gene optimization for expression of Crimean-Congo haemorrhagic fever nucleoprotein

Introduction: Crimean-Congo haemorrhagic fever (CCHF) virus causes a severe and often fatal infection. The virus has the propensity to cause nosocomial infections and hence a rapid and sensitive diagnosis is important for isolation of the patient for protection of health care workers and implementation of supportive therapy. Current serological diagnostic assays are based on enzyme linked immunosorbent assays (ELISA) or immunofluorescence (IF) tests using inactivated virus which necessitates biosafety level (BSL) 4 facilities for preparation of reagents. The aim of the study was to prepare a safe recombinant nucleoprotein (NP) and determine its suitability for detection of antibody responses in survivors of CCHF infection. To facilitate protein expression the gene encoding the nucleoprotein was optimized for *E.coli* usage.

Materials and Methods: The CCHF nucleoprotein gene coding sequence was submitted to GenScript (USA). OptimumGene software was used to optimize codon usage, GC content, eliminate polyadenylation sites and modify cis-acting sites. The optimized gene was synthesized and the nucleoprotein expressed in an *E.coli* bacterial expression system with a His tag for purification. CCHF antibody positive human sera were tested in an ELISA using the recombinant NP antigen.

Results: A direct ELISA was developed in which plates were coated with the recombinant NP antigen and reacted with convalescent human sera from confirmed CCHF patients. The antigen expressed from the optimized NP gene was found to detect IgG antibodies against CCHF virus.

Conclusion: Recombinant proteins have been shown to be safe, cost effective reagents that can be prepared for biohazardous pathogens without the requirements of BSL 4 facilities. Optimization of the CCHF N gene was essential for high expression of soluble NP. The protein expressed was shown to react against antibodies in convalescent CCHF patients. A panel of serum samples from confirmed patients will be used to determine the specificity and sensitivity of the ELISA for diagnostic purposes.

APPENDIX 6

Title and abstract of presentation at the Faculty of Health Sciences AstraZeneca, Research Forum 27-28th August 2010, University of the Free State, Bloemfontein

Gene optimization for expression of Crimean-Congo haemorrhagic fever virus nucleoprotein and induction of humoral immunity in mice

Awarded trophy for best junior laboratory paper presented at the health Sciences AstraZeneca Faculty Research Forum 2010

Authors: Samudzi R, Swanepoel R, Leman P, Paweska J, Burt FJ.

Department of Medical Microbiology and Virology, Special Pathogens Unit, National Institute for Communicable Disease, Johannesburg

Presenter: Samudzi R

Crimean-Congo haemorrhagic fever virus (CCHFV) causes a severe and often fatal infection. The virus has the propensity to cause nosocomial infections hence a rapid diagnosis is important. Current serological diagnostic assays are based on reagents using inactivated virus which necessitates biosafety level (BSL) 4 facilities for preparation. This study aims at using gene optimization to improve protein expression of the nucleoprotein of CCHFV in *E.coli* and to determine if the expressed protein is functional in an immunoassay and immunogenic in an animal model.

Materials and Methods: The CCHFV NP gene coding sequence was submitted to GenScript. OptimumGene software was used to optimize codon usage, GC content, eliminate polyadenylation sites. The optimized gene was synthesized and the nucleoprotein expressed in an *E.coli* bacterial expression system. Confirmed CCHF antibody positive human sera were tested in an ELISA using the recombinant NP antigen. To determine the immunogenicity of the recombinant CCHFNP the protein was inoculated into mice in combination with adjuvant.

Results: A direct ELISA was developed in which plates were coated with the recombinant NP antigen and reacted with human sera from confirmed CCHF patients. The antigen expressed from the optimized NP gene detected IgG antibodies against CCHFV in all patient sera. The recombinant NP induces a detectable humoral antibody response in serum samples collected from immunized mice using immunofluorescent tests.

Conclusion: Recombinant proteins are safe, cost effective reagents that can be prepared for biohazardous pathogens without the requirements of BSL 4 facilities. Codon optimization of the CCHF N gene was essential for high expression of NP. The protein expressed was shown to react against antibodies in convalescent CCHF patients and to induce a detectable humoral antibody response in mice.