Design, expression, and purification of a Flaviviridae polymerase using a high-throughput approach to facilitate crystal structure determination

Kyung H. Choi, James M. Groarke, Dorothy C. Young, Michael G. Rossmann, Daniel C. Pevear, Richard J. Kuhn and Janet L. Smith

Protein Sci. 2004 13: 2685-2692
Access the most recent version at doi:10.1110/ps.04872204

Design, expression, and purification of a *Flaviviridae* polymerase using a high-throughput approach to facilitate crystal structure determination

KYUNG H. CHOI,[1] JAMES M. GROARKE,[2,3] DOROTHY C. YOUNG,[2] MICHAEL G. ROSSMANN,[1] DANIEL C. PEVEAR,[2] RICHARD J. KUHN,[1] AND JANET L. SMITH[1]

[1]Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-2054, USA
[2]Department of Biology, ViroPharma Inc., Exton, Pennsylvania 19341, USA

(RECEIVED May 18, 2004; FINAL REVISION May 18, 2004; ACCEPTED June 20, 2004)

Abstract

Bovine viral diarrhea virus (BVDV) nonstructural protein 5B is an RNA-dependent RNA polymerase, essential for viral replication. Initial attempts to crystallize a soluble form of the 695-residue BVDV polymerase did not produce any crystals. Limited proteolysis, homology modeling, and mutagenesis data were used to aid the design of polymerase constructs that might crystallize more readily. Limited proteolysis of the polymerase with trypsin identified a domain boundary within the protein. Homology modeling of the polymerase, based on the structure of hepatitis C virus polymerase, indicated that the two polymerases share a 23% identical "core," although overall sequence identity is low. Eighty-four expression clones of the BVDV polymerase were designed by fine-sampling of chain termini at the boundaries of domain and of active truncated forms of the polymerase. The resulting constructs were expressed in *Escherichia coli* and purified using high-throughput methods. Soluble truncated proteins were subjected to crystallization trials in a 96-well format, and two of these proteins were successfully crystallized.

Keywords: bovine viral diarrhea virus; hepatitis C virus; *Flaviviridae*; NS5B; RNA-dependent RNA polymerase; high-throughput; limited proteolysis

Bovine viral diarrhea virus (BVDV), a single-stranded RNA virus, is a member of the pestivirus genus in the *Flaviviridae* family. The *Flaviviridae* family also includes human pathogens, such as dengue virus, West Nile virus, and hepatitis C virus (HCV). BVDV causes a wide variety of clinical disease syndromes in cattle, including infertility, abortion, and congenital defects in calves (Fray et al. 2000). Outbreaks of these syndromes cost the U.S. cattle industry more than a billion dollars per year, and BVDV continues to be the most devastating and economically important viral disease of cattle (Wyovet 2003). Despite the availability of vaccines, the most effective way of controlling the disease remains the elimination of persistently infected animals, which are the main carriers of the virus (Wyovet 2003).

About 3% of the world population is chronically infected with HCV, which causes significant liver disease including chronic hepatitis, cirrhosis, and hepatocellular carcinoma (World Health Organization 1999; Penin et al. 2004). Because HCV does not grow well in cell culture systems and BVDV gene expression and translation strategy is similar to that used by HCV, BVDV has been used as a model system for HCV (Buckwold et al. 2003).

Reprint requests to: Janet L. Smith, Department of Biological Sciences, Purdue University, 915 West State St., West Lafayette, IN 47907-2054, USA; e-mail: smithj@purdue.edu; fax: (765) 496-1189.

[3]Present address: Novartis Institutes for Biomedical Research, Cambridge, MA 02139, USA.

Abbreviations and symbols: BDV, border disease virus; BVDV, bovine viral diarrhea virus; CSFV, classical swine fever virus; HCV, hepatitis C virus; HT, high-throughput; IPTG, isopropyl-β-D-thiogalactoside; MDV, mucosal disease virus; NS, nonstructural; PDB, Protein Data Bank; PVDF, polyvinylidene fluoride; RdRp, RNA-dependent RNA polymerase; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.04872204.

BVDV has a positive-sense RNA genome of 12.6 kb that is translated as a single polyprotein. The polyprotein is processed by viral and host cellular proteases into at least four structural (C, E1, E2, E^{rns}) and six nonstructural (NS) proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) required for viral assembly and replication (Meyers and Thiel 1996; Lindenbach and Rice 2001). Among NS proteins, the functions of NS3, NS4A, and NS5B have been characterized. NS3 contains an N-terminal viral serine protease domain, which requires NS4A as a co-factor, and a C-terminal helicase/NTPase domain. NS5B, the largest NS protein, is an RNA-dependent RNA polymerase (RdRp) of 719 amino acids. The membrane-bound replication complex, which contains the RNA polymerase, the helicase, and other viral and cellular proteins, is essential for viral genome replication (Lindenbach and Rice 2001). For this reason, the polymerase is an attractive target for antiviral drug development. Several inhibitors preventing the replication of BVDV have been characterized (Baginski et al. 2000; Sun et al. 2003). Viruses with reduced susceptibility to these inhibitors have been isolated and shown to have mutations in the RNA polymerase coding region. The structure of BVDV polymerase–drug complexes will facilitate elucidation of the mechanism of drug inhibition and is essential for structure-based drug design. However, initial attempts to determine the structure of the BVDV polymerase were hampered by poor solubility of the protein and subsequent difficulties producing crystals. BVDV polymerase presents challenging crystallization problems: (1) multiple domains connected by flexible hinges, (2) a putative membrane-associated region, and (3) an additional domain of unknown (probably non-polymerase) function. The truncation of the potential membrane-associated region (C-terminal 24 amino acids) of the polymerase (Lai et al. 1999) did not result in crystal production. Dynamic light scattering experiments indicated that the truncated protein aggregated even in high salt concentration. The goal for crystallographic studies is to provide a structure from which antiviral drugs can be designed. Thus, a detailed view of the polymerase active site is essential.

Recently, high-throughput (HT) cloning, expression, and protein purification systems have been developed for structural genomics. In this approach, multiple open reading frames from a genome are separately cloned and expressed without prior knowledge of the structure or function of the encoded proteins. The three-dimensional structures of these proteins are then determined using various techniques, primarily X-ray crystallography. The protocols used in the HT system can be adapted for parallel production of multiple constructs of a single protein target. The clones can be screened systematically for soluble, active proteins amenable to crystallization. However, it is difficult to predict which protein modifications will influence protein solubility and, thus, crystallization. This approach allows all available knowledge about the target protein, such as function and domain boundaries, to be incorporated into rational design of constructs.

In the present study, limited proteolysis and homology modeling based on HCV polymerase were employed to aid in the design of BVDV polymerase constructs. These constructs were cloned, expressed, and proteins purified using an HT approach. About half of the constructs produced soluble proteins suitable for structural and functional studies. Two constructs yielded crystals that were useful for X-ray crystallographic investigations. This method is well-suited to a robotic approach and is applicable to other problematic proteins.

Results

Multiple sequence alignment and homology model

A homology model of BVDV polymerase assisted the design of constructs for expression and crystallization. Among polymerases of known structure, the only example from the *Flaviviridae* family and nearest relative to BVDV polymerase is the enzyme from HCV (Ago et al. 1999; Bressanelli et al. 1999; Lesburg et al. 1999). However, the sequence identity of BVDV and HCV polymerases is quite low (13%). Thus, critical evaluation of a multiple sequence alignment was a prerequisite to homology modeling. The homology modeling procedure consisted of three steps: First, multiple sequence alignment of BVDV polymerase with the polymerases of three other pestiviruses and HCV was used to identify conserved residues that are likely to be important for polymerase structure and function. Second, the conserved residues were mapped onto the three-dimensional structure of the HCV polymerase to identify a "core" structure common to both BVDV and HCV polymerases. Third, the core domain identified in the second step was used to design constructs that eliminate possible loops and domains that may not be essential for RNA polymerase function.

The sequences of the polymerases of the pestiviruses BVDV, classical swine fever virus (CSFV), border disease virus (BDV), and mucosal disease virus (MDV) were aligned simultaneously with the HCV polymerase sequence. The initial alignment was performed using the CLUSTAL W program (Thompson et al. 1994), followed by manual alignment of eight RdRp sequence motifs of positive-sense RNA viruses (Koonin 1991). The pestivirus polymerase amino acid sequences are 70%–75% identical. All pestivirus polymerases are about 130 amino acids longer at the N terminus than HCV polymerase, and these additional residues are likely to form a distinct structural domain. The overall sequence identity between pestivirus and HCV polymerases is less than 15%. However, the central core of the protein, residues 248–527 of BVDV polymerase, is 23%

identical to the corresponding residues of HCV polymerase (residues 126–396). This core region includes eight sequence motifs characteristic of RdRps of positive-sense RNA viruses (Fig. 1A).

A homology model of the core region of BVDV polymerase was made by mapping the core sequence determined by multiple sequence alignment onto the three-dimensional structure of HCV polymerase (Protein Data Bank [PDB] accession code 1GX5; Fig. 1B). Polymerase structures have been described in terms of a right hand comprised of "fingers," "palm," and "thumb" domains. The palm domain is the catalytic domain, which contains five RdRp motifs, including an essential GDD motif. The thumb domain is most variable among polymerases. The homology model constructed from residues 126 to 396 of HCV polymerase corresponds to part of the fingers domain (residues 126–175 and 228–279), the entire palm domain (residues 176–227 and 280–372), and a few residues of the thumb domain (residues 387–396; Fig. 1B). The thumb domain of HCV polymerase (residues 397–530) has no detectable sequence identity with the C-terminal region of BVDV polymerase. Thus, BVDV polymerase is likely to have similar fingers and palm domains to those of HCV polymerase, but a different thumb domain and an additional domain at its N terminus.

Limited proteolysis

Proteases with a variety of cleavage specificities were tested to identify compact domains of BVDV polymerase. Chymotrypsin and thermolysin were tested based on their broad specificity, and trypsin was tested for its cleavage preference at basic residues since BVDV polymerase is a basic protein in which 10% of all residues are Lys or Arg. Only trypsin cleavage of BVDV polymerase (lacking the C-terminal 24 amino acids) yielded stable fragments when analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Two protein fragments of 64 and 55 kDa were stable up to one hour after incubation with 2% (w/w) trypsin (Fig. 2). N-terminal sequencing identified Arg132–Asn133 as the cleavage site for both fragments. The calculated molecular weight from residue 133 to the C terminus is 64 kDa. Thus, the 55 kDa band is the product of a second cleavage, approximately 80 amino acids from the C terminus. The trypsin cleavage at Arg 132 is consistent with the multiple sequence alignment, which suggested that BVDV polymerase has a distinct N-terminal domain of ~130 amino acids, and with mutational studies of the polymerase, in which up to 90 amino acids were deleted from the N terminus without the loss of polymerase activity (Lai et al. 1999).

Based on limited proteolysis and homology modeling of BVDV polymerase, 84 truncation constructs were designed for HT cloning and protein production (Fig. 3). Crystalliza-

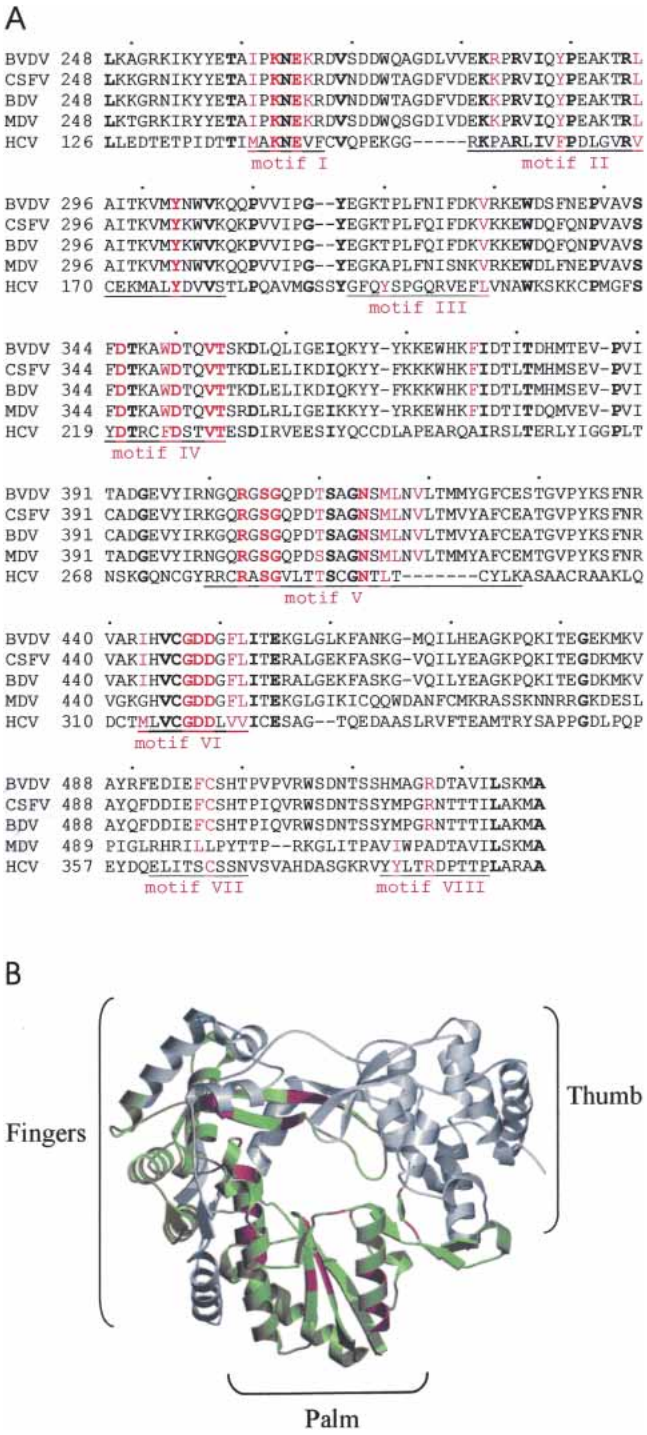

Figure 1. Multiple sequence alignment and homology model of BVDV polymerase. (*A*) Sequence comparison of the common core region of pestivirus and HCV polymerases. Pestiviruses include BVDV, bovine viral diarrhea virus; CSFV, classical swine fever virus; BDV, border disease virus; and MDV, mucosal disease virus. The conserved residues are shown in bold. The RdRp motifs of positive-sense RNA viruses are labeled and colored in red. (*B*) Structure of HCV polymerase (PDB accession code 1GX5). The core residues, which share higher sequence identity between pestivirus and HCV polymerases, are shown in green, and the RdRp motifs are red. All other residues are shown in gray.

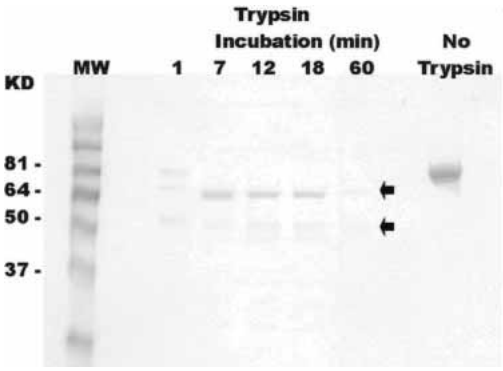

Figure 2. Limited proteolysis of BVDV polymerase with trypsin. Two fragments indicated by arrows were sequenced at the N terminus, and the cleavage site was identified as being between Arg132 and Asn133 for both fragments.

tion behavior is often sensitive to the state of polypeptide chain termini, particularly nonnative termini. Thus, 13 N-terminal start sites spanning residues 62–145 in steps of 5–10 amino acids were investigated, including the trypsin cleavage site at residue 133. Smaller N-terminal truncations at positions 62, 71, 79, 83, and 89 represented constructs shown to be active in previous experiments (Lai et al. 1999). Five C-terminal truncations at residues 678, 683, 687, 691, and 695 were designed based on the tryptic digest results indicating C-terminal cleavage. Disorder of the C-terminal thumb domain has been reported in crystal structures of some polymerases (Jeruzalmi and Steitz 1998); thus, two C-terminal truncations at residue 500 and 540 were designed to remove the thumb domain entirely. The expression constructs included combinations of the N- and C-terminal truncations with either N- or C-terminal hexahistidine tags (His-tags; Fig. 3).

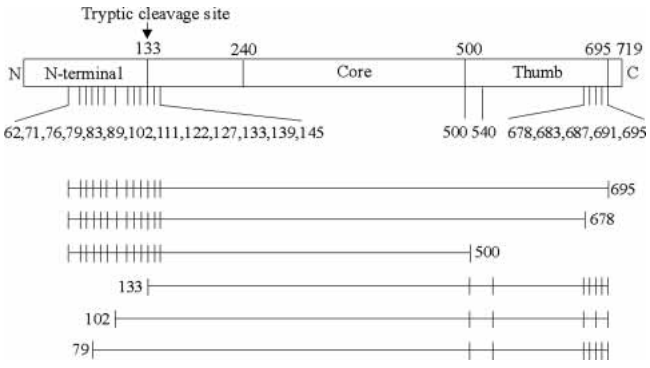

Figure 3. Design of BVDV polymerase mutants. The overall structure of BVDV polymerase can be divided into an N-terminal domain, a core domain that contains a part of the "fingers" domain and "palm" domain, and a "thumb" domain. Vertical bars *below* the schematic represent truncation sites utilized in this study. For each set of constructs represented by horizontal lines, all combinations of N and C termini were made. Most of the proteins were expressed twice, with either an N-terminal or C-terminal His-tag. Some of the constructs were made as a single His-tagged protein.

HT cloning and protein expression of BVDV polymerase mutants

The overall strategy of HT expression/purification and the number of successful clones at each step are summarized in Figure 4 and Table 1, respectively. The HT process, from PCR amplification to expression, was performed in a week. The 84 PCR primers were designed to have similar melting temperatures, between 50°–60°C, because all reactions were run in one 96-well plate. PCR amplification resulted in 79 constructs with positive bands as analyzed by agarose gel electrophoresis (Fig. 5). After restriction enzyme digestion and ligation, *Escherichia coli* strain BL21 (DE3) RIL was transformed directly with the ligation mixture without purification of the restriction enzyme digest. Three colonies of each protein construct were picked and grown in 96-deep-well plates. SDS-PAGE analysis showed that 48 clones were expressed and their total cell extracts yielded proteins of the expected molecular weight. Additionally, XL10-Gold competent cells were transformed with the ligation mixture for plasmid preparation for later use.

The desired protein constructs were quickly and efficiently identified by Western dot blot analysis of cell extracts (Fig. 6). Among the 48 expressing constructs, 45 showed a positive reaction in Western blot analysis with BVDV polymerase antibodies. The soluble fractions of the 45 positive clones were further analyzed by SDS-PAGE.

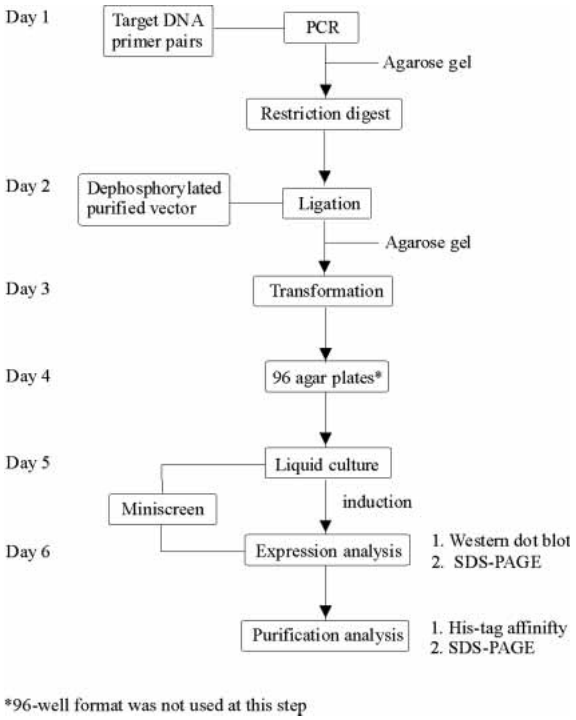

Figure 4. Diagram of HT cloning and expression. All steps were performed in 96-well format except for the step labeled with an asterisk (*).

Table 1. *Summary of HT cloning and protein purification*

	Number of successful clones	Success rate (%)
Number of mutants designed	84	
PCR positive	79	94
Cloning positive	48	61
Expression-positive dot blot	45	94
Soluble proteins (SDS-PAGE)	40	89
His-tag binding and elution	22	55
Purified proteins	15	
Crystallized proteins	2	

Five constructs, which were positive in the dot blot assay, did not show protein bands of the calculated molecular weight and were not included in the next purification step.

HT purification

Forty proteins of the correct molecular weight were tested for exposed His affinity tags, which would be used for large-scale purification by affinity chromatography. Thirty-one protein constructs bound to a Co-chelating resin in 96-well plate format; however, only 22 protein constructs eluted with 150 mM imidazole, a concentration that is much higher than that used to purify the full-length form of BVDV polymerase (30–50 mM; Choi et al. 2004). Of the 22 proteins that were eluted from the Co-chelating resin on a small scale, 15 were expressed in 1-L cultures and purified for crystallization trials. The proteins purified on a large scale were BVDV 79–500, 79–678, 79–683, 102–687, 102–695, 111–695, 127–695, 133–500, and 139–500 with an N-terminal His-tag, and BVDV 71–679, 79–500, 79–540, 111–679, 122–695, and 133–500 with a C-terminal His-tag.

Crystallization

The construct comprising residues 71–679 of the BVDV polymerase (BVDV 71–679) produced tiny crystals (<20 μm) within one week in the presence of 2.0 M ammonium sulfate and either 0.1 M sodium acetate (pH 4.6), or 0.1 M phosphate-citrate (pH 4.2). Crystallization conditions were screened with different concentrations of ammonium sulfate and a series of buffers from pH 4.0 to pH 6.0 at ~0.3 unit increments. Two constructs, BVDV 71–679 and BVDV 79–678, produced crystals in one month from 5% isopropanol and 2.0 M ammonium sulfate. Improved crystals were obtained by screening varying ammonium sulfate concentrations in combination with small alcohols, such as methanol, ethanol, 1-propanol, butanols, or ethylene glycol. The best crystals were obtained from 1.6 M ammonium sulfate and either 4% isopropanol or 15% ethylene glycol. When other purified proteins were screened against the above conditions, none produced crystals except BVDV 79–683, which produced tiny crystals after long (>2 months) incubation at 20°C. Thus, only the proteins with 678, 679, or 683 as a C terminus crystallized. The polymerase sequence between 685–695 contains patches of hydrophobic residues, which may have prevented crystallization in addition to the flexibility of the C terminus. Constructs of the polymerase lacking the thumb domain (proteins with either residue 500 or 540 as a C terminus) were soluble, but were unstable, degraded within days of purification, and did not result in any crystals. Removal of the His-tag was not investigated because good-quality crystals grew from tagged protein. Interestingly, BVDV 71–679 has a C-terminal His-tag and BVDV 79–678 has an N-terminal His-tag. The three-dimensional structures of BVDV 71–679 and 79–678 have now been determined by X-ray crystallography (Choi et al. 2004).

Discussion

Soluble protein is a prerequisite for crystal growth and subsequent structure determination by X-ray crystallography. Although great improvements have been made in being able to use less protein for a full structure determination, production of stable, soluble protein remains a major bottleneck for many important problems. Several strategies have been described for overcoming protein solubility and crystallization problems. A common and often successful approach is to crystallize a similar protein from an alternative biological source. Often, however, a specific protein is targeted for crystallization, which may need to be modified to achieve solubility. Gene fusion with a large affinity tag can be useful for increasing protein expression and solubility (Smyth et al. 2003). Several fusion proteins, including thioredoxin, maltose binding protein, glutathione S-transferase, intein, and calmodulin-binding protein, have been used to generate soluble proteins (Shih et al. 2002; Smyth et al. 2003). Often, however, the fused tag prevents crystallization due to conformational heterogeneity of a flexible

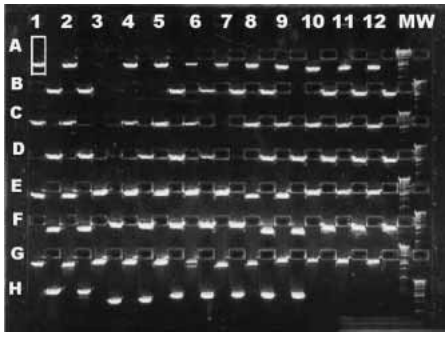

Figure 5. PCR amplification. PCR products were loaded onto an E-gel (1% agarose) in a 96-well format with a molecular weight marker and visualized with ethidium bromide. Lane *A1* is outlined for clarity.

linker peptide, which must be removed in an additional purification step. Engineering proteins with single-residue substitutions can also increase solubility (Wingren et al. 2003), although many substitutions must be tested to achieve the desired solubility. Recently, a directed evolution approach has been used to produce soluble proteins when wild-type protein is insoluble (Pédelacq et al. 2002; Yang et al. 2003). In the case of multidomain proteins, truncation of the protein to smaller functional domains can be effective. Functional domains can be determined by limited proteolysis, multiple sequence alignment, or homology modeling, as in the present study. Limited proteolysis has been used to trim loose ends of a protein for crystallization, making the protein more compact and conformationally homogeneous and, thus, more likely to crystallize than the full-length protein (Matsui et al. 2004). However, due to the potential problem of increasing heterogeneity, limited proteolysis is not commonly used. Recombinant technology is preferred to generate fragments identified by proteolysis, but without the heterogeneity that accompanies proteolysis. If soluble protein is obtained, but does not yield crystals, fine tuning of N- and C-termini of the protein may be required to produce or improve crystals. In the HT approach outlined here, generation of stable domains having variable N- or C-termini was performed in parallel, vastly reducing time and resource requirements. This approach is particularly suited to problems in which one has no estimate, a priori, of which constructs are likely to be most successful at the expression, purification, or crystallization steps.

The strategy used to express soluble and crystallizable forms of the BVDV polymerase can be readily adapted to other problematic proteins. Multiple sequence alignment is a useful tool for identifying conserved domains even in cases where there is no three-dimensional structure of a homologous protein available. For limited proteolysis, several proteases, in combination with variable digestion conditions, can be tested to determine the best conditions for each protein. Information and previous experience regarding the target protein should be considered when designing protein constructs.

The HT procedure used in this study was performed manually in a 96-well format, but can be transferred easily to robotic liquid handling systems. Every step in the HT process can be automated, including steps in which the 96-well format was not used, such as colony picking. The restriction enzyme digestion or ligation step can be replaced with the ligation-independent cloning system. Ninety-six-well SDS-PAGE gels are now available, which further reduces time and labor.

The identification of clones producing soluble proteins is one of the crucial steps in the HT process. Antibodies against BVDV polymerase were used to detect the expressed protein in this study, but fusion tag antibodies can be used to identify proteins, when a specific antibody is not available (Knaust and Nordlund 2001). For the purification of many constructs, it is practical to screen an affinity purification step on a small scale. Metal affinity chromatography, crystal screening kits, and crystallization plates are well suited for the 96-well format.

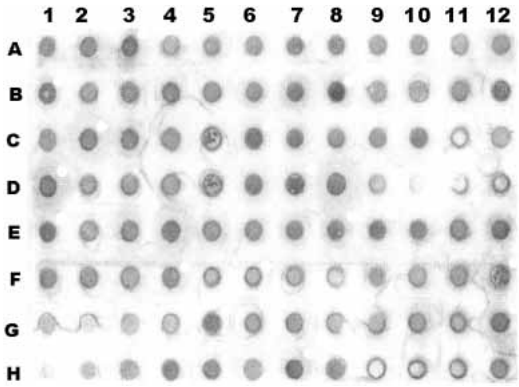

Figure 6. Protein expression analysis using Western dot blot. After cell lysis, the supernatant was transferred to PVDF membrane and protein expression was monitored by Western dot blot analysis using BVDV polymerase antibodies. Wells H1 and H7 represent negative and positive controls, respectively.

The crystal structures of BVDV polymerases (BVDV 71–679, BVDV 79–678) have been determined to 2.9 Å resolution (Choi et al. 2004). The polymerase possesses the well known right-hand shape, with an independent N-terminal domain that HCV polymerase lacks. In these BVDV polymerase constructs, residues 71 to 139 form an N-terminal domain of which residues 71–90 are disordered. The trypsin cleavage site at Arg132 is found in a loop connecting the N-terminal domain to the fingers domain. The fingers and palm domains of BVDV polymerase are similar to those in HCV polymerase, but the thumb domain is significantly different. The BVDV polymerase structure is consistent with the multiple sequence alignment and the limited proteolysis results, confirming their power for probing the structure of proteins. A similar approach has been extended to other viral polymerases and has resulted in the successful production of soluble proteins in our laboratories.

Materials and methods

Multiple sequence alignment

Pestivirus polymerases, which include BVDV (GenBank accession code P19711), CSFV (NP_777506), BDV (AF037405), MDV (NC_001461), and HCV polymerase (P29846), were aligned using the CLUSTAL W program (Thompson et al. 1994). Additional manual adjustment of the alignment was necessary to align the known RdRp sequence motifs of positive-sense RNA viruses (Koonin 1991).

The core region, which shares moderate sequence identity between pestivirus and HCV polymerases, was mapped onto the three-dimensional structure of HCV polymerase (PDB accession code 1GX5) using the program O (Jones et al. 1991).

Limited proteolysis

Full-length BVDV polymerase lacking 24 C-terminal amino acids (0.1 mg) was incubated with 2% (w/w) trypsin (2 μg) (Sigma) in 0.1 M NH_4HCO_3 (pH 8.2) at room temperature. Aliquots (10 μl) were taken at several time points, quenched with SDS-containing sample buffer, and analyzed by SDS-PAGE. The desired bands in SDS-PAGE gel were transferred to a polyvinylidene fluoride (PVDF) membrane using a Trans-Blot electrophoretic transfer cell (BioRad) for 1 h at 4°C. N-terminal amino acid sequencing was performed at the Protein Separation and Analysis Laboratory of Purdue University.

PCR amplification

Primers were generated to encode either an N-terminal His-tag containing a thrombin cleavage site or a C-terminal His-tag for expression using plasmids pET28 and pET30 (Novagen), respectively. For N-terminal His-tagged proteins, the 5′-primers were designed to have an AseI restriction site and the 3′-primers to have the stop codon (TGA) followed by an XhoI restriction site. C-terminal His-tagged proteins were generated with NdeI (or AseI) and XhoI restriction sites. The 5′-primers were designed to have an AseI restriction site (ATT AAT) followed by G to generate the start codon (ATG) and 3′-primers with an XhoI restriction site (CTC GAG). Primer sequences were adjusted such that each oligonucleotide had a calculated melting temperature of 50°–60°C.

Plasmid encoding BVDV polymerase (BVDV NADL strain) lacking 24 C-terminal amino acids was used for all subsequent manipulations (Baginski et al. 2000). PCR amplication was performed with the Expand High Fidelity PCR system (Roche). PCR products as well as positive controls were loaded onto an E-gel 96 HT agarose electrophoresis system (1% agarose gel, Invitrogen) and visualized by ethidium bromide staining. Amplified PCR products were purified with the QIAquick 96 PCR fragment purification kit (Qiagen Inc.).

Restriction-enzyme digest of PCR products and vectors

The purified PCR products were digested simultaneously with restriction enzymes AseI and XhoI. The enzymes were added to 96-well plates containing the purified PCR products and incubated at 37°C for 2 h. The pET28 and pET30 vectors were digested with NdeI and XhoI. After the restriction enzyme digestion, the vectors were dephosphorylated using shrimp alkaline phosphatase (Roche) and gel purified before use.

Ligation and transformation

The appropriate vectors and Ready-to-go T4 DNA ligase (Amersham Pharmacia Biotech) in water were added to the PCR products in 96-well plates. The ligation plate was incubated overnight at 16°C, and aliquots from each well were loaded onto the E-gel 96 HT agarose electrophoresis system. The ligation products were transformed into either XL10-Gold competent cells or BL21-CodonPlus RIL competent cells (Stratagene). Transformants of BL21 cells were grown overnight on agar plates with 34 μg/ml kanamycin and 30 μg/ml chloramphenicol. Three colonies per construct were picked manually for growth in liquid culture. *E. coli* cells were grown at 37°C in 1.3 ml of Luria broth with 34 μg/ml kanamycin and 30 μg/ml chloramphenicol to an optical density (OD_{600}) of 0.5–0.7 in a 96-deep-well plate. Protein expression was induced by 1.5 mM isopropyl-β-D-thiogalactoside (IPTG) at 18°C overnight.

Western dot blot analysis

The cells for each construct in triplicate were transferred to the corresponding wells of a polystyrene 96-well plate and centrifuged at 3000 rpm for 10 min. The supernatant was removed and 250 μl of lysis buffer (extraction buffer, 1 tablet of Complete Protease Inhibitor [Roche], 35 μl of β-mercaptoethanol per 100 ml of lysis buffer) was added directly to each well. Cells were lysed using a Misonix Sonicator 3000 (Misonix). After centrifugation, 100 μl of supernatant was transferred to a PVDF membrane using a dot blot vacuum unit. Anti-rabbit BVDV polymerase antibodies (Collett et al. 1988) and goat anti-rabbit horseradish peroxidase (Sigma) were used as primary and secondary antibodies, respectively. Proteins were visualized with 3,3'-diaminobenzidine 3,3',4,4'-tetraaminobiphenyl 4 HCl developing solution. The positive clones from the Western blot were further analyzed by SDS gel electrophoresis.

Metal affinity purification

Each 96-well plate was filled with one SwellGel cobalt chelated disc (Pierce). Supernatant (200 μl) of the lysed cells was transferred directly to the resin and left for 5 min at 4°C. The plate was centrifuged at 1500 rpm for 10 min. The flow-through was saved, and the resins were washed five times with wash buffer (50 mM sodium phosphate at pH 7.0 and 0.5 M NaCl) to remove unbound proteins followed by centrifugation at 1000 rpm for 10 min. His-tagged proteins were eluted with the wash buffer containing 5, 50, and 150 mM imidazole and analyzed by SDS-PAGE.

Large scale purification

One liter of *E. coli* cell cultures were grown in Luria broth containing 34 μg/ml kanamycin and 30 μg/ml chloramphenicol at 37°C to an optical density (OD_{600}) of 0.7. Protein expression was induced with 1.5 mM IPTG and the cultures grown overnight at 18°C. The cells were pelleted and lysed in a French press using 50 ml of extraction buffer (CLONTECH) supplemented with 0.5 M NaCl, 5 mM β-mercaptoethanol, and one tablet of Complete Protease Inhibitor (Roche). The lysate was centrifuged at 15,000 rpm for 10 min, and the supernatant was mixed with 5 ml of Talon resin (CLONTECH) at 4°C for 30 min. The resin was washed five times with 20 ml of wash buffer A (CLONTECH extraction/wash buffer supplemented with 1 M NaCl and 1% Nonidet P-40) and five times with 20 ml of wash buffer B (CLONTECH extraction/wash buffer supplemented with 0.5 M NaCl). The Talon resin was then transferred to a column, and weakly bound proteins were removed by washing with 50 ml of wash buffer B containing 5 mM imidazole. His-tagged BVDV polymerase was eluted from the Talon resin with buffer B containing 30 mM and 50 mM imidazole in a stepwise manner, and fractions were analyzed by SDS-PAGE. Protein-containing fractions were pooled and dialyzed against 20 mM Tris-Cl (pH 7.5), 0.3 M NaCl, 5% glycerol, and 2 mM DTT for 16 h at 4°C. The dialyzed protein was mixed with DEAE-sepharose resin and incubated at 4°C for 30 min. The resin was centrifuged and the supernatant applied to a Poros-HS (Perceptive) cation exchange column pre-equilibrated with 20 mM Tris-Cl (pH 7.5), 0.3 M NaCl, 5% glycerol, and 2 mM DTT. The protein was eluted

with a linear gradient of 0.3 to 0.6 M NaCl. Most of the proteins eluted between 0.4 and 0.5 M NaCl. Final purity of the proteins was >95% based on SDS-PAGE. Typical yields were ~10–20 mg of purified polymerase protein per liter of bacterial culture.

Crystallization

Purified proteins were concentrated to ~10 mg/ml in 50 mM Tris-Cl (pH 7.5), 0.4 M NaCl, and 5% glycerol using an ultrafree centrifugal filter device (Millipore). Initial crystallization conditions were screened using vapor-diffusion sitting-drop geometry in 96-well plates with Hampton and Emerald Biostructure crystallization screening kits.

Acknowledgments

We thank Christopher T. Jones and Todd Geders for their help in HT cloning; and Cheryl Towell and Sharon Wilder for help in the preparation of the manuscript. The work benefited from a Purdue University reinvestment fund and was supported by an NIH Program Project Grant (AI 55672) to R.J.K., J.L.S., and M.G.R. K.H.C. was supported by the Purdue University Trask fund.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K., and Miyano, M. 1999. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure* **7:** 1417–1426.

Baginski, S.C., Pevear, D.C., Seipel, M., Sun, S.C.C., Benetatos, C.A., Chunduru, S.K., Rice, C.M., and Collett, M.S. 2000. Mechanism of action of a pestivirus antiviral compound. *Proc. Natl. Acad. Sci.* **97:** 7981–7986.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R.L., Mathieu, M., De Francesco, R., and Rey, F.A. 1999. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc. Natl. Acad. Sci.* **96:** 13034–13039.

Buckwold, V.E., Beer, B.E., and Donis, R.O. 2003. Bovine viral diarrhea virus as a surrogate model of hepatitis C virus for the evaluation of antiviral agents. *Antiviral Res.* **60:** 1–15.

Choi, K.H., Groarke, J.M., Young, D.C., Kuhn, R.J., Smith, J.L., and Rossmann, M.G. 2004. The structure of the RNA-dependent RNA polymerase from bovine viral diarrhea virus establishes the GTP binding site required for de novo initiation. *Proc. Natl. Acad. Sci.* **101:** 4425–4430.

Collett, M.S., Larson, R., Belzer, S.K., and Retzel, E. 1988. Proteins encoded by bovine viral diarrhea virus: The genomic organization of a pestivirus. *Virology* **165:** 200–208.

Fray, M.D., Paton, D.J., and Alenius, S. 2000. The effects of bovine viral diarrhoea virus on cattle reproduction in relation to disease control. *Anim. Reprod. Sci.* **60–61:** 615–627.

Jeruzalmi, D. and Steitz, T.A. 1998. Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. *EMBO J.* **17:** 4101–4113.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47:** 110–119.

Knaust, R.K.C. and Nordlund, P. 2001. Screening for soluble expression of recombinant proteins in a 96-well format. *Anal. Biochem.* **297:** 79–85.

Koonin, E.V. 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* **72:** 2197–2206.

Lai, V.C.H., Kao, C.C., Ferrari, E., Park, J., Uss, A.S., Wright-Minogue, J., Hong, Z., and Lau, J.Y.N. 1999. Mutational analysis of bovine viral diarrhea virus RNA-dependent RNA polymerase. *J. Virol.* **73:** 10129–10136.

Lesburg, C.A., Cable, M.B., Ferrari, E., Hong, Z., Mannarino, A., and Weber, P.C. 1999. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat. Struct. Biol.* **6:** 937–942.

Lindenbach, B.D. and Rice, C.M. 2001. *Flaviviridae*: The viruses and their replication. In *Fields virology*, 4th ed. (eds. D.M. Knipe and P.M. Howley), pp. 991–1041. Lippincott Williams & Wilkins, Philadelphia.

Matsui, T., Hogetsu, K., Akao, Y., Tanaka, M., Sato, T., Kumasaki, T., and Tanaka, N. 2004. Crystallization and X-ray analysis of the N-terminal core domain of a tumour-associated human DEAD-box RNA helicase, rck/p54. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 156–159.

Meyers, G. and Thiel, H.J. 1996. Molecular characterization of pestiviruses. *Adv. Virus Res.* **47:** 53–118.

Pédelacq, J.D., Piltch, E., Liong, E.C., Brerendzen, J., Kim, C.Y., Rho, B.S., Park, M.S., Terwilliger, T.C., and Waldo, G.S. 2002. Engineering soluble proteins for structural genomics. *Nat. Biotechnol.* **20:** 927–932.

Penin, F., Dubuisson, J., Rey, F.A., Moradpour, D., and Pawlotski, J.M. 2004. Structural biology of hepatitis C virus. *Hepatology* **39:** 5–19.

Shih, Y.-P., Kung, W.-M., Chen, J.-C., Yeh, C.-H., Wang, A.H.-J., and Wang, T.-F. 2002. High-throughput screening of soluble recombinant proteins. *Protein Sci.* **11:** 1714–1719.

Smyth, D.R., Mrozkiewicz, M.K., McGrath, W.J., Listwan, P., and Kobe, B. 2003. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **12:** 1313–1322.

Sun, J.-H., Lemm, J.A., O'Boyle II, D.R., Racela, J., Colonno, R.J., and Gao, M. 2003. Specific inhibition of bovine viral diarrhea virus replicase. *J. Virol.* **77:** 6753–6760.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.

Wingren, C., Edmundson, A.B., and Borrebaeck, C.A.K. 2003. Designing proteins to crystallize through-strand pairing. *Protein Eng.* **16:** 255–264.

World Health Organization. 1999. Global surveillance and control of hepatitis C. Report of a WHO consultation organized in collaboration with the viral hepatitis prevention board. *J. Viral Hepat.* **6:** 35–47.

Wyovet, Wyoming Animal Health and Disease Information Network, Wyoming State Veterinary Laboratory. 2003. BVD in Wyoming update. http://www.wyovet.uwyo.edu.

Yang, J.K., Park, M.S., Waldo, G.S., and Suh, S.W. 2003. Directed evolution approach to a structural genomics project: Rv2002 from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci.* **100:** 455–460.