A Conserved Tripeptide Sequence at the C Terminus of the Poxvirus DNA Processivity Factor D4 Is Essential for Protein Integrity and Function*

Received for publication, October 7, 2016, and in revised form, November 11, 2016 Published, JBC Papers in Press, November 11, 2016, DOI 10.1074/jbc.M116.761908

Manunya Nuth§, Hancheng Guan‡, and Robert P. Ricciardi‡§§

From the §Department of Microbiology, School of Dental Medicine and the ¶Abramson Cancer Center, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Edited by Charles E. Samuel

Vaccinia virus (VACV) is a poxvirus, and the VACV D4 protein serves both as a uracil-DNA glycosylase and as an essential component required for processive DNA synthesis. The VACV A20 protein has no known catalytic function itself but associates with D4 to form the D4-A20 heterodimer that functions as the poxvirus DNA processivity factor. The heterodimer enables the DNA polymerase to efficiently synthesize extended strands of DNA. Upon characterizing the interaction between D4 and A20, we observed that the C terminus of D4 is susceptible to perturbation. Further analysis demonstrated that a conserved hexapeptide stretch at the extreme C terminus of D4 is essential for maintaining protein integrity, as assessed by its requirement for the production of soluble recombinant protein that is functional in processive DNA synthesis. From the known crystal structures of D4, the C-terminal hexapeptide is shown to make intramolecular contact with residues spanning the inner core of the protein. Our mutational analysis revealed that a tripeptide motif (G215F217) within the hexapeptide comprises apparent residues necessary for the contact. Prediction of protein disorder identified the hexapeptide and several regions upstream of Gly215 that comprise residues of the interface surfaces of the D4-A20 heterodimer. Our study suggests that 215GFI217 anchors these potentially dynamic upstream regions of the protein to maintain protein integrity. Unlike uracil-DNA glycosylases from diverse sources, where the C termini are disordered and do not form comparable intramolecular contacts, this feature may be unique to orthopoxviruses.

Processivity factors serve to tether their cognate DNA polymerases onto the template to permit the synthesis of long extended strands of DNA. In the absence of its processivity factor, the polymerase will fall off the template after limited nucleotide incorporation. Processivity factors are categorized into two classes. The first class consists of ring-shaped assemblies of identical subunits that completely encircle the DNA template through the action of ATP-driven clamp loaders (1, 2). Well studied processivity factors that are members of this class include the Escherichia coli DNA polymerase III β subunit (3, 4) and eukaryotic proliferating cell nuclear antigen (5, 6). The second class consists of processivity factors that do not completely encircle the DNA template and, hence, do not require clamp loaders. The members of this class include the herpesviruses, as represented by cytomegalovirus, Epstein-Barr virus, herpes simplex virus type 1, human herpesvirus 6, and Kaposis sarcoma-associated herpesvirus (7–13).

Poxviruses are large (130–230 kbp), enveloped, double-stranded DNA viruses that replicate exclusively in the cytoplasm. Viral DNA synthesis is dependent upon at least six virally encoded proteins (14). Vaccinia virus (VACV) is the prototypic poxvirus sharing strong DNA sequence homology with other orthopoxviruses, of which variola, the causative agent of smallpox, is a member. Although nearly all DNA viruses encode a single protein that functions as a processivity factor, poxviruses are unique in that the processivity factor is composed of two virally encoded proteins: A20 and D4 (15–22).

A20 is a 49-kDa protein with no identified catalytic activity, but it can associate with D4 and viral DNA polymerase (18, 19, 21). D4 is a 25-kDa protein that is essential for viral replication and viability (15–17). Because D4 is unable to bind viral DNA polymerase directly, it utilizes A20 as a bridge. In this manner, D4 indirectly tethers viral DNA polymerase onto the DNA template. With its inherent uracil-DNA glycosylase (UDG) activity, it was recently hypothesized that the template-scanning action of the UDG is utilized by D4 to confer the processive DNA synthesis activity by the DNA polymerase (22).

Despite exhibiting poor protein sequence homology, D4 shares similar structural core features with UDGs of other organisms (23, 24). As evidenced by X-ray crystallography, isolated D4 is homodimeric (24), whereas biochemical evidence supports the formation of the D4-A20 heterodimer (21, 22, 25). Recent structural evidence further supports the heterodimer model when both A20 and D4 are present (26, 27). This would

* This work was supported by National Institutes of Health Grants 5U01-A1082211 and 1R44AI115759-01. The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

† To whom correspondence should be addressed: Dept. of Microbiology, School of Dental Medicine, University of Pennsylvania, 221 Levy Research Bldg., 4010 Locust St., Philadelphia, PA 19104. Tel.: 215-898-3905; E-mail: ricciardi@upenn.edu.

§ The abbreviations used are: VACV, vaccinia virus; UDG, uracil-DNA glycosylase; SPR, surface plasmon resonance; MBP, maltose binding protein; ITC, isothermal titration calorimetry; MCV, molluscum contagiosum virus; TEV, tobacco etch virus; Ni-NTA, nickel-nitrilotriacetic acid; HSQC, heteronuclear single-quantum coherence.
suggest the formation of the heterodimer as energetically favored and, therefore, a likely biologically relevant model of the processivity factor. With respect to the protein binding surfaces on D4, the crystal structures of the D4 homodimer and the D4-A20 heterodimer show overlapping residues at the protein-protein interface (24, 27). Both of the contact surfaces are relatively large and featureless and involve a significant amount of hydrophobic residues, with the preference for the heterodimer suggested by the larger contact surface area (1890 Å² of buried surface) and key hydrogen bonds not observed (27) for the D4 homodimer (1510 Å² of buried surface) (24).

To gain a greater understanding of the formation of the D4-A20 heterodimer, we sought to characterize the biophysical interaction of D4 with A20 by surface plasmon resonance (SPR). The results of our binding studies indicated complex protein binding behavior. The reported crystal structures of D4 (either as a homodimer or heterodimer) implied that it is relatively rigid. This was a surprising revelation that provided the impetus to investigate its solution properties further. Ultimately, deletion analysis showed that a short hexapeptide stretch at the C terminus of D4 is necessary for maintaining protein integrity and function. From the known crystal structures of D4, it appears that the hexapeptide makes contact with residues spanning the core of the protein. Using mutagenesis studies, we further revealed that a tripeptide motif (215GFI217, based on VACV numbering) within the hexapeptide is necessary for maintaining the intramolecular contact. We discuss how this unique intramolecular contact serves to maintain protein integrity and function.

**Results**

**SPR Reveals That the Binding of A20 Is Dependent on the Presentation of D4**—The D4-A20 heterodimer functions as the poxvirus processivity factor required for viral DNA synthesis (21, 22, 25–27). Structural support for the heterodimer has been demonstrated (26, 27). In the case of A20, the full-length protein has not been successfully isolated to investigate its biophysical properties, such as structure elucidation and interaction with D4. Thus, although the production of full-length A20 remains a challenge, the truncated versions of this protein have provided valuable insights. Indeed, the use of a yeast two-hybrid assay had earlier identified the N-terminal 25 amino acids of A20 as the minimal requirement for D4 binding, with increased affinity observed when extended to 50 amino acids and longer (25). The *de novo* protein folds of both the N-terminal 25 and 50 amino acids were predicted to consist of two α helices (28), with the two helical structures of the N-terminal 50 amino acids demonstrated by X-ray crystallography to be in complex with D4 (27).

On the basis of the yeast two-hybrid findings (25), an N-terminal 63-amino acid truncation of A20 was chosen for this study. Prediction showed that an extra α helix is introduced, compared with the 50-amino acid extension (28), that may further contribute to protein-protein interactions. Fusing the N-terminal 63 amino acids to the maltose binding protein (MBP-A20₁₋₆₃) permitted the expression of sufficient quantities of soluble product (Fig. 1A). Significantly, the requirement of MBP as a carrier protein was demonstrated by the inability to express soluble N-terminal regions of A20 up to 100 amino acids in the absence of MBP in both bacteria and baculovirus (Fig. 1, B and C).
A Tripeptide of D4 is Essential for Processive DNA Synthesis

We initially interrogated the binding of D4 to A20 by isothermal titration calorimetry (ITC) because the solved crystal structure and studies by column chromatography (27) suggested that the formation of the heterodimer is facile and, therefore, amenable to this label-free approach. However, we routinely observed protein precipitation during ITC experiments (data not shown). Therefore, an adsorption method, such as SPR, was chosen as an effective alternative to ITC. As the ligand, D4 was prepared with either an N- or C-terminal His6 tag to permit capture onto a CM5 biosensor chip cross-linked with anti-His antibody. The rationale was that, when the D4-A20 heterodimer assembles, aggregates that might otherwise serve as seeds for protein precipitation would be incapable of forming because of the separation of nearby heterodimers by virtue of attachment to the cross-linked antibody under a constant buffer flow. The N- or C-terminal tagged protein, His6D4 or D4His6, respectively, was captured onto the active cell using dilute protein preparations (2 μg/ml) in the presence of 25% (w/v) glycerol. The binding experiments were then performed in running buffer lacking glycerol and with added BSA to minimize nonspecific interactions. As demonstrated by the negligible response to captured His6D4, the injections of up to 20 μM purified MBP served as a negative control (Fig. 2A). By comparison, there was pronounced protein binding with the injections of up to 5 μM MBP-A201–63 (Fig. 2A). Specifically, we found improved curve fitting when constrained to the local response Rmax as opposed to global Rmax. This was likely due to non-uniform ligand captures between replicates. The observed 3-fold increase in affinity for captured D4His6 (KD = 0.21 μM) was imparted by the apparent slower dissociation of the D4His6-A20 heterodimer (Fig. 2). In a reciprocal experiment, MBP-A201–63 was captured by cross-linked anti-MBP antibody and injected with D4 proteins. In this case, the sensograms showed non-first-order behavior and, therefore, did not permit data fitting (data not shown), which would be supportive of heterogeneity of D4 proteins. This indicated that D4 should be used as the ligand in subsequent experiments. Notably, these SPR results demonstrated that the binding of A20 to D4 varied by the positioning of the His6 tag.

**Protein Binding Thermodynamics**—Next, we sought to understand the D4-A20 interaction in greater detail by examining its binding thermodynamics by SPR. When measured up to 35 °C, it was noticed that the sensograms for both captured D4 protein versions showed a poor fit (data not shown), an attribute commonly associated with gross structural change. Therefore, data collections from 10–30 °C were used. For both captured D4 versions, a linear van’t Hoff fit was established at 10–25 °C, giving rise to favorable entropy (TΔS = 11.93 and 7.24 kcal/mol for His6D4 and D4His6, respectively), whereas nonlinearity was demonstrated when analyzed up to 30 °C (Fig. 2, C and E).

**Assessment of Protein Structural Integrity and Impact of the His6 Tag**—To ensure that the observed difference in the SPR studies was not due to disruption of the protein fold by the tag placement, D4 proteins were examined by CD spectroscopy. As shown in Fig. 3A, the spectra of all three protein versions showed comparable far-UV traces with extracted secondary structure contents in agreement with the reported crystal structures (24). D4His6 in particular showed a slight decrease in near-UV absorbance, suggesting a potential environment of the aromatic residues to be different from that of either untagged D4 or His6D4 (Fig. 3A).

The protein environment was further examined by 1H-15N heteronuclear single-quantum coherence (HSQC) NMR spectroscopy. Because untagged D4 was generated from the same 15N-labeled His6D4 preparation, this permitted a reasonable comparison of the D4 proteins differing only in the tag status. It is worth mentioning that 15N-labeled D4His6 was not amenable to NMR studies because of low protein expression in minimal medium and the tendency to precipitate when concentrated, thus providing anecdotal evidence of protein instability when the tag was introduced at the C terminus. Unlike the solved crystal structures (24, 27), both untagged D4 and His6D4 pro-

---

3 M. Nuth, unpublished observation.
A Tripeptide of D4 is Essential for Processive DNA Synthesis

**FIGURE 3.** A–D, the impact of the His<sub>6</sub> tag on the structural and stability properties of D4 proteins as determined by far- and near-UV CD spectroscopy (A), <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectroscopy (B and C), and steady-state UV-visible spectroscopy (D). D, the values for the soluble fractions were determined relative to untreated proteins. All values reflect the mean ± S.D. of three repeats. The <i>p</i> values were determined by analysis of variance in comparison with the untagged D4 for the corresponding temperature treatments. *, <i>p</i> < 0.001; rt, room temperature; MRW, mean residual weight.

proteins showed fewer than expected numbers of amide HN cross-peaks for a 218-amino acid protein and exhibited poor dispersion at <sup>1</sup>H chemical shifts (6.5–7.5 ppm; Fig. 3, B and C). Significantly, the impact of the tag was demonstrated by the reduction of peaks for His<sub>6</sub>D4 (Fig. 3C) compared with untagged D4 (Fig. 3B). Overall, although the CD results showed no evidence of disruption of the protein fold by either tag placement, the NMR studies suggested a difference in protein environment between the tagged and untagged proteins.

To corroborate the NMR results, the impact of the tag on protein stability was further examined by measuring soluble protein fractions after 2-h incubation at room temperature (30, 37, and 42 °C). In general, protein instability was observed with the introduction of the His<sub>6</sub> tag. For example, ~50% of untagged D4 protein precipitated at 37 °C, whereas <10% of His<sub>6</sub>D4 and D4His<sub>6</sub> remained soluble (Fig. 3D). Specifically, the impact of the C-terminal His<sub>6</sub> tag was evident from the significant protein precipitation at 30 and 37 °C compared with either His<sub>6</sub>D4 or untagged D4 (Fig. 3D). At 42 °C, all protein versions failed to stay in solution (Fig. 3D). Taken together, the CD results showed that the His<sub>6</sub> tag at either the N or C terminus did not disrupt the protein fold, thus providing support for the hypothesis that the proteins used in the SPR experiments were not structurally impaired. However, the NMR and stability studies consistently demonstrated the impact of the His<sub>6</sub> tag on protein properties, with the C terminus seemingly susceptible to perturbation on the basis of the stability results.

Sequence Conservation among the C TERMINI OF ORTHOPOXVIRUSES—Prior to revealing its essential role in processive DNA synthesis, VACV D4 was discovered to harbor UDG activity (15–17). An interesting hallmark of UDGs is their remarkable structural conservation despite exhibiting poor overall protein sequence homology (29, 30). Nevertheless, a short stretch near the C terminus comprised of amino acids <sup>209</sup>Pi(N/D)<sup>W</sup>212 (based on VACV numbering) is conserved for poxviruses, <i>E. coli</i>, <i>Homo sapiens</i>, and HSV-1 (Fig. 6A). This feature was observed earlier (23) but with unknown significance. Unique to the orthopoxviruses is a highly conserved stretch of six amino acids (<sup>213</sup>AQGFIY<sup>218</sup>) based on VACV numbering) downstream of the <sup>209</sup>Pi<sup>W</sup>212 sequence, with minor variations observed for raccoonpox and Yoka poxvirus (Fig. 6A). In the crystal structures, a feature seemingly unique to VACV (and likely applicable to other Orthopoxvirus members) is the presence of short secondary structure elements: an α helix generated by <sup>211</sup>NWAQ<sup>214</sup> and a β sheet comprised of <sup>216</sup>FI<sup>217</sup> (24). By comparison, the crystal structures of the UDG examples showed random coils at the C terminus (PDB codes are given in Fig. 6A).

The Extreme C Terminus of D4 Is Required for Protein Integrity and Function—Because our biophysical results pointed to the susceptibility of the C terminus of D4 to perturbation, we next sought to establish whether this region of the protein was indeed important for maintaining integrity and function. Therefore, we focused on the conserved hexapeptide stretch <sup>213</sup>AQGFIY<sup>218</sup>. Starting from Tyr<sup>218</sup>, each amino acid was sequentially deleted, and the detection of soluble protein fractions by Western blotting analysis of the bacterial preparations served as an assessment of protein integrity (Fig. 4, A and B). The sequential removal of amino acids from Ile<sup>217</sup> to Ala<sup>213</sup> resulted in D4 proteins fractionating exclusively as inclusion bodies, whereas the deletion of the terminal Tyr<sup>218</sup> retained solubility (Fig. 4B). Protein misfolding was considered the most likely cause leading to the formation of inclusion bodies.

We next investigated the functional importance of this region of D4 by expressing these D4 deletion mutants in a cell-free system for use in the measurement of processive DNA synthesis activity. As shown in the autoradiogram in Fig. 4C, the protein expression levels for all D4 mutants were comparable with the wild type. Notably, unlike the bacterial preparations, all eukaryotic translated D4 mutant proteins were detected in
the soluble fractions. We speculated that this was due to the binding and retention of the misfolded proteins to the abundant soluble molecular chaperones present in the reticulocyte lysate (31) used in the cell-free system. By contrast, the significant increase in protein expression levels, in combination with potentially incompatible chaperone system, may have promoted the inclusion body formation observed in the bacterial preparations. Processive DNA synthesis was then assessed by combining D4 proteins with similarly expressed proteins of A20 and viral DNA polymerase as before (21, 32). It is significant to note that this cell-free method permitted the expression of full-length VACV DNA polymerase and full-length A20 (33).

The corresponding D4 protein levels are shown by autoradiography of [35S]cysteine incorporation. To permit sample order, the image of the WT is a splice of the crude lysate of the uninduced cells. WB, Western blot. C, in vitro DNA synthesis activity was measured using proteins produced by the cell-free method.

Within the Hexapeptide Stretch, 215GFF217 Is Indispensable for Protein Integrity and Function—Interestingly, the crystal structure of D4 shows the hexapeptide making intramolecular contact with regions of the inner core of the protein (Fig. 7A), with 213AQ214 adopting elements of an α helix and 215GF217 a β sheet (24). By employing mutagenesis, we sought to identify residues within the hexapeptide that may be essential for maintaining the intramolecular contact. Initially, alanine substitutions (with the exception of Ala213) were performed. Substitutions of Gly215 and Phe216 caused D4 to fractionate exclusively into the inclusion bodies, whereas I217A produced both insoluble and minor soluble products (Fig. 5B). By comparison, alanine mutants of the flanking residues (Gln214 and Tyr218) had no effect on protein solubility (Fig. 5B). Each of these mutants was then tested for processive DNA synthesis activity. Only G215A and F216A were unable to support processive DNA synthesis (Fig. 5D), which was in accordance with the loss of protein integrity.

Because the results of the alanine mutations suggested the importance of Gly215 and Phe216 for maintaining protein integrity and function, additional substitutions of the hexapeptide were explored. Therefore, the expression profiles of the various point mutations were examined. Proteins were scored as soluble, insoluble, or in between. The in between designation represented the relatively smaller fraction of soluble proteins detected by Western blotting analysis, e.g. I217A (Fig. 5B). As summarized in Fig. 6B, only alterations at the Gly215 and Phe216 positions were not tolerated, as demonstrated by the exclusive production of inclusion bodies. Specifically, the conservative F216Y mutation yielded reduced levels of soluble protein, indicating tolerability of only a minor alteration at this position (Fig. 6B). In accordance with the results of Y218D, the introduction of either alanine or phenylalanine at Tyr218 had no effect on protein solubility (Fig. 6B). Taken together, the 215GFF217 motif, particularly 215GF216, was critical for maintaining protein integrity.

Discussion

Recent structural studies of D4 in complex with A20 (27, 34) have provided insightful details into the formation of this unique heterodimeric processivity factor necessary for poxvirus DNA synthesis. Although the D4-A20 heterodimer is regarded as the relevant form of the processivity factor (21, 22, 25–27), it is notable that the heterodimer and the D4 homodimer share overlapping residues at the protein interface. The crystal structure suggests increased surface interactions as
A Tripeptide of D4 is Essential for Processive DNA Synthesis

FIGURE 6. Properties of the C-terminal region of the D4 protein. A, protein sequence alignment of the C terminus of VACV D4 to representative UDGs. Shown are representatives from orthopoxviruses (CMLV, camelpox; CPX, cowpox; ECTV, mousepox; HSPV, horsepox; MPV, monkeypox; RCNV, raccoonpox; TPV, tanapox; VARV; and YKPV, Yoka poxvirus), cervipoxvirus (DPV, deerpox), avipoxvirus (FWPV, fowlpox), capripoxvirus (GTPV, goatpox), molluscipoxvirus (MCV), leporipoxvirus (MVX, myxoma virus), parapoxvirus (ORFV, orf virus), suipoxvirus (SWPV, swinepox), yatapoxvirus (YLD, Yaba-like disease virus), and UDGs with reported crystal structures (Ec, E. coli; Hs, H. sapiens). Of note, only VACV is structurally available among the orthopoxviruses. The maroon solid line marks the hexapeptide stretch. B, protein solubility summary of single amino acid substitutions at the hexapeptide stretch. Proteins from bacterial expressions that were undetectable by Western blotting are classified as insoluble (red), whereas detections in both the crude and soluble fractions are identified as soluble (green). Proteins detected with a minor presence in the soluble fraction are identified as in between (orange). Efforts were made to choose substitutions that would alter the physical and/or chemical properties at the sites of interest. Therefore, mutations that either dramatically interfered with protein solubility with only minor changes (e.g. the Gly215 site) or produced no effects with significant substitutions (e.g. the Gln214 site) were pursued only sparingly. Unpursued mutations are shown uncolored. The asterisk indicates that the assessment is from the C-terminal swap of MCV, which contains Val217 in addition to218PL219 (Fig. 5A).

FIGURE 5. A–D, protein expression and function of the C-terminal swaps (A and C) and alanine mutants (B and D). Conditions and notations are the same as in Fig. 4. C and D, all values reflect the mean ± S.D. of three repeats, with the p values shown in comparison with the WT. D, the autoradiographic image of the WT is a splice of the same gel to depict sample order. WB, Western blot; S, soluble fraction; C, crude lysate; U, crude lysate of uninduced cells; Ec, E. coli; Hs, H. sapiens.
A Tripeptide of D4 is Essential for Processive DNA Synthesis

a possible reason for the preference for the heterodimer (27). Therefore, to form the heterodimer, A20 must displace the D4-D4 interaction, presumably with favorable energetics gained from these increased surface interactions. However, we found that the use of subdenaturing guanidinium HCL (≤1 M), which is thought to contribute to the structural and energetic stabilization of proteins (35), and the decreased solvent dielectric upon introduction of dilute ethanol (1%-3%) was incapable of fractionating D4 into monomers (data not shown). Rather, protein aggregation was observed (data not shown), suggesting that D4 is sensitive to perturbation. In this work, we reported the first biophysical characterization of the interaction between D4 and A20 and identified complex protein behavior exerted by D4.

The SPR results demonstrated A20-binding characteristics to be impacted by ligand presentation of D4, thus indicating protein orientation (and, hence, protein conformational constraint) on the biosensor chip as an important consideration. Both the C- and N-terminally labeled proteins showed similar binding thermodynamic features, as evidenced by the favorable entropy, which highlights the importance of hydrophobic contributions (22, 27), and an associated conformational change demonstrated by the ensuing negative heat capacity change. Moreover, the use of MBP as a carrier protein proved to be necessary for A20 protein production. Given that A20 cannot stay soluble without MBP, the observed crystal structure of the N-terminal 50-amino acid (27) suggests that A20 is a disordered/unfolded or partially folded cargo that adopts a protein fold upon binding to D4. Therefore, the orientation of D4 on the biosensor chip surface could conceivably affect this folding event and give rise to the observed difference in affinity between the two tagged versions. Notably, submicromolar affinities were determined for both captured D4 versions. Because of the likelihood of a disordered/unfolded or partially folded state of A20, conformational heterogeneity may have contributed to rebinding events that led to slower $k_{off}$ rates that were largely responsible for the affinity determination. This is consistent with the enhanced data dispersion observed for sensorgrams of both captured D4 versions when A20 was used at concentrations in excess of 5 μM (Fig. 2, A and B). Therefore, the reported $K_D$ values may not accurately reflect the assumed simple bimolecular interaction between D4 and A20. Importantly, the findings revealed the sensitivity of D4 to tag placement and reiterated the sensitivity of the protein to perturbation. These experimental cues prompted us to investigate its solution properties further.

Overall, the CD results showed no ill effects of the tag on protein fold. The NMR spectra, however, revealed fewer than expected $^1$H-$^1$H HSQC peaks and regions of poor dispersion for both tagged and untagged proteins. These are surprising findings because the crystal structures reported well folded proteins (24, 27). Therefore, our data lend credence to the aggregation tendency of D4 as opposed to the disruption to protein folding by the introduced tag. Indeed, the impact of the tag was further corroborated in the promotion of protein sizing heterogeneity observed by dynamic light scattering (data not shown), whereby the increased protein-protein interactions underlying protein aggregation can be supported by the significant reduction of $^1$H-$^1$H HSQC peaks when the tag was present.

Consistently, protein instability was more pronounced for the C-terminally tagged protein compared with either the N-terminally tagged or untagged protein. Its destabilizing effect is exemplified by the striking loss of ~80% of soluble protein after 2-h incubation at 30 °C (Fig. 3D). Taken together, D4 is shown to be susceptible to perturbation, and the introduction of a His$_6$ tag at the C terminus is even more disruptive than either the N-terminally tagged or untagged protein. Therefore, the tendency to dimerize (or aggregate) is likely due to the compensation for incurred unfavorable energetics should D4 become monomeric. Indeed, no structures of the monomeric D4 are reported.

The crystal structures of D4, either as a homodimer (24) or in complex with the A20 peptide (27) or DNA (34, 36), depict D4 as relatively rigid. By comparison, conformational dynamics are shown to be important to confer DNA site recognition by UDGs (37, 38). Because our findings infer the monomeric form of D4 as an energetic liability, we sought to investigate protein disorder as a likely explanation because it plays important roles in molecular recognition and protein function (39, 40). As shown in Fig. 7B, five protein-spanning regions are predicted to be disordered by DisEMBL according to the “hot loops” definition (41). Notably, the D4 interface surfaces comprising the D4 homodimer (24) and D4-A20 heterodimer (27) are accounted for within the predicted C-terminal regions 165–191 and 201–218, whereas regions 69–85 and 165–191 contain sites shown to contact DNA (34, 36) (Fig. 7B). Importantly, the hexapeptide falls within the predicted disordered region 201–218 (Fig. 7B). Given that disordered proteins can become ordered when bound to other molecules (42, 43), it follows that the existence of secondary features within the hexapeptide stretch (24) is due to intramolecular interaction. Therefore, our mutagenesis results strongly support the $^{215}$GF$^{217}$ motif as key residues responsible for the contact.

As summarized in Fig. 7C, it is speculated that the existence of disordered regions upstream of the hexapeptide, as an energetic liability, favors the formation of the complexed forms of D4. In the presence of A20, the heterodimer is formed, whereas the homodimer is formed in the absence of A20. This likely explains why monomeric D4 is never isolated. For instance, the crystal structures of the D4 homodimer (PDB codes 2OWQ and 2OWR) contain regions of missing coordinates (24) that are suggestive of such disorder. Therefore, it follows that the intramolecular contact by 215-GFI-217 helps to anchor these presumably disordered/dynamic regions that are important for partner recognition (e.g. A20 and DNA). In addition to the orthopoxviruses, other genera of poxviruses, interestingly, share the indispensable $^{215}$GF$^{216}$ motif (Fig. 6A), thus suggesting a conserved mechanism. These findings, in total, clearly demonstrate that the removal or alteration of the $^{215}$GF$^{217}$ residues on the extreme C terminus leads to the loss of protein integrity and function.

Experimental Procedures

Materials—Carbodiimide hydrochloride, N-hydroxysuccinimide, ethanolamine, and CMS biosensor chips were purchased
A Tripeptide of D4 is Essential for Processive DNA Synthesis

FIGURE 7. Protein disorder and the role of $^{215}\text{GFI}^{217}$. A, the diagram depicts D4 as a monomer (PDB code 20WR), with the hexapeptide shown as a green ribbon and the side chains of $^{215}\text{GFI}^{217}$ as orange sticks. The hexapeptide sequence is shown with the corresponding color scheme. In the crystal structure, residues spanning the region 102–113 (magenta) are in close proximity to $^{215}\text{GFI}^{217}$ and are presumably responsible for the bulk of the intramolecular contact.

B, the five regions of disorder predicted by DisEMBL. C, a proposed model depicting the role of $^{215}\text{GFI}^{217}$ in maintaining protein integrity and, thus, function. Through intramolecular interaction, $^{215}\text{GFI}^{217}$ (shown with a similar color scheme as in A) anchors the potentially dynamic regions of D4 (with only the predicted region 201–218 shown highlighted in red) that are important for target recognition (e.g. A20 and DNA). The removal of $^{215}\text{GFI}^{217}$ or the weakening of the intramolecular contact by amino acid substitutions untethers the dynamic region of D4 and is an energetic liability. As a result, protein misfolding ensues, leading to protein heterogeneity and aggregation and, ultimately, protein precipitation. Because of the hydrophobic nature of the protein interface, it is presumed that the binding surfaces of D4 will be buried within these aggregates. In this scenario, A20 or DNA is incapable of binding to soluble aggregated D4 proteins because of the lack of access to the binding surfaces. As an explanation for the observed instability exhibited by D4 in solution, the promotion of protein aggregation and heterogeneity is likely rooted in dynamics still present for the homodimer.

Cloning, Mutagenesis, and Protein Expression—A20R and D4R genes were PCR-amplified from genomic DNA of VACV, WR strain (a generous gift from Profs. G. H. Cohen and R. Eisenberg). The cloning of D4R with either the NdeI/BamHI or NcoI/XhoI sites into the pET28b(+) vector (Novagen) afforded the expression of the N- or C-terminal His$_6$-tagged protein version, respectively. Additionally, the thrombin cleavage site of the vector (downstream of the N-terminal His$_6$) was mutated to a TEV protease cleavage site (corresponding to the protein sequence ENLYFQS) by site-directed mutagenesis to permit efficient removal of the tag. Full-length A20R was introduced into pET15b(+) (Novagen) at the NdeI/BamHI sites, and stop codons were introduced by site-directed mutagenesis to generate the N-terminal 50-, 63-, and 100-amino acid truncated proteins. The corresponding N-terminal 63-amino acid region of A20R was inserted into pMAL-c2x (New England Biolabs) via the EcoRI/XhoI sites to encode a 63-amino acid fusion protein (MBP-A20$_{1–63}$). The generation of MBP alone was accomplished by introducing a stop codon at the EcoRI site of pMAL-c2x. For baculovirus expression, the N-terminal 100-amino acids of A20R was cloned into pBacPAK8 (Clontech Laboratories) at the XhoI/EcoRI sites with an introduced N-terminal His$_6$ tag. All sequences were confirmed by DNA sequencing.

Protein expression was performed using the E. coli strain Rosetta 2(DE3)LyS5 (Novagen). For expression of the D4 constructs, fresh LB broth, supplemented with 50 µg/ml kanamycin and 30 µg/ml chloramphenicol, was inoculated with a 1:200 volume of overnight cell cultures, grown to an A$_{600}$ of 1 at 37°C, and induced overnight with 0.2 mM isopropyl β-D-1-thiogalactopyranoside at 20°C. A20 (constructs of pET15b(+)), MBP, and MBP-A20$_{1–63}$ were similarly cultured in LB broth supplemented with 100 µg/ml ampicillin and 30 µg/ml chloramphenicol. For baculovirus expression, SF9 cells were cultured in a spinner tube to ~2 × 10$^6$ cells/ml in 100 ml of SF-900 II medium (Life Technologies) supplemented with 100 units/ml of penicillin and 100 µg/ml of streptomycin and infected with A20 recombinant virus at a 0.4 multiplicity of infection. Cells were harvested at 50% cytopathic effect as determined by trypan blue staining.

Cells expressing D4 constructs were lysed in Ni-NTA column buffer (50 mM sodium phosphate (pH 8), 0.4 M NaCl, and 0.6 M sucrose) containing 0.5% Triton X-100, 5 mM benzamidine, and 1 mM phenylmethylsulfonyl fluoride by freeze-thaw and repeated sonication. For examination of the soluble and insoluble protein fractions, cell pellets collected from 3-ml growths were lysed in 500 µl of the above buffer, and 60 µl of crude or cleared lysate was diluted into 200 µl of 1× SDS-PAGE gel loading buffer. 10- and 5-µl samples were used for Coomasie staining and Western blotting analysis, respectively. For purification, the lysate was clarified by centrifugation at 15,000 rpm for 30 min, adjusted with 10 mM imidazole, and loaded onto an Ni-NTA column. The column was subsequently washed with 25 volumes of 80 mM imidazole and eluted with 5 volumes of 200 mM imidazole. The eluent was concentrated and further purified by gel filtration (Superose 200, GE Healthcare) equilibrated with 20 mM sodium phosphate (pH 7), 150 mM NaCl, and 1 mM EDTA (working buffer). Cells harboring MBP and MBP-A20$_{1–63}$ were similarly lysed in amylene buffer (30 mM sodium phosphate (pH 7.8), 150 mM NaCl, and 1 mM EDTA) supplemented with 5 mM benzamidine and 1 mM phenylmethylsulfonyl fluoride, clarified, bound to amylene resins, washed with 25 volumes of amylene buffer, and eluted with five volumes of 15 mM maltose. The eluent was then concentrated and similarly purified by gel filtration. Proteins were confirmed by Western blotting using anti-His (anti-tetraHis, mouse
A Tripeptide of D4 is Essential for Processive DNA Synthesis

The Journal of Biological Chemistry

Monoclonal antibody, catalog no. 34670, lot no. 145022663, Qiagen) or anti-MBP (mouse monoclonal antibody, catalog no. E8032S, lot no. 0091305, New England Biolabs) primary antibody for the detection of D4 or MBP and MBP-A20_{1–63}, respectively, and the purity was judged by SDS-PAGE. Protein concentrations were determined by A_{280} using the predicted extinction coefficients 45,630, 84,800, and 66,350 cm^{-1} M^{-1} for D4 constructs, MBP-A20_{1–63}, and MBP, respectively (45).

**TEV Protease Digestion**—The protein fractions from the Ni-NTA column elutions were added to 20 mM β-mercaptoethanol, 1 mM EDTA, and 1:50 TEV protease and digested by 2-day incubation at 4 °C. After two rounds of dialysis into 10 volumes of Ni-NTA column buffer at 4 °C, the dialysate was passed through an Ni-NTA column twice, and the flow-through, containing only the digested D4, was collected and finally purified by gel filtration. The procedure also removed the TEV protease, which harbored a His_{6} tag. The removal of the N-terminal His_{6} tag was confirmed by SDS-PAGE and the lack of protein capture onto an anti-His surface on a biosensor chip (data not shown). The resulting untagged D4 is predicted to contain a serine upstream of the first methionine.

**Protein Handling**—Because of the tendency of D4 to precipitate during the freeze-thaw cycle, proteins used for all studies were freshly prepared, kept at 4 °C, and used within 2 days after the last purification step.

**SPR**—Experiments were performed on a Biacore X100 (GE Healthcare) using filtered and degassed buffers. Both flow cells of a CM5 biosensor chip were activated for 7 min at 5 μl/min with a 1:1 mixture of 391 and 100 mM carbodiimide hydrochloride and N-hydroxysuccinimide, respectively, in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, and 0.005% P20. Anti-tetraHis antibody was desalted prior to use, prepared at 5 μg/ml in 10 mM sodium acetate (pH 5), and injected at 5 μl/min for 15 min to achieve ~4000 and 2500 response units onto the reference and active cells, respectively, after quenching for 15 min with 0.5 mM ethanolamine. Ligand capture onto the active cell was accomplished by injecting 2 μg/ml His_{6}D4 or D4His_{6} diluted in 20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 1 mM EDTA, 0.2 mg/ml carboxymethyl dextran, and 0.005% P20 (running buffer) containing 25% (w/v) glycerol. Approximately 50–80 response units were captured after a 360 s injection at 5 μl/min. Analytes (MBP or MBP–A20_{1–63}) were prepared in running buffer at 0, 0.125, 1.25, 2.5, 5, 10, and 20 μM concentrations and added to 1 mg/ml BSA. The temperature-dependent experiments were recorded at 10, 15, 20, 30, and 35 °C, and MBP–A20_{1–63} was prepared at 0, 1.25, 2.5, and 5 μM concentrations. For a typical binding experiment, and immediately after surface capture and a 30-s buffer wash, the analyte was injected for 180 s at 30 μl/min, monitored for 180 s, and regenerated for 20 s with 50% (v/v) ethylene glycol in 50 mM glycine/50 mM phosphate (pH 2.5). The active surface was then recaptured and repeated with different analyte concentrations. The 0 μM in each cycle was used for double referencing. All curves were fit to a 1:1 Langmuir with local R_{max} by the included BIAevaluation software, yielding chi-square values of ≤3.4 for all experiments. The temperature dependence of binding was examined by linear fit from 10–25 °C and nonlinear fit to the integrated form of the van’t Hoff equation (46) for 10–30 °C.

**UV-visible and CD Spectroscopies**—For the determination of protein solubility, 20 μM protein was diluted in working buffer in a 50-μl volume and incubated at the indicated temperature for 2 h. Particulates were removed by centrifugation at 14,000 rpm for 5 min, and the absorbance of the supernatant was recorded at 280 nm. Experiments were repeated three times. CD measurements were recorded at 25 °C on an Aviv Model 410 spectrometer (Aviv Biomedical) of ~5 and ~20 μM proteins prepared in working buffer for far-UV (190–260 nm) and near-UV (260–320 nm) determinations, respectively. Measurements were recorded at 1-nm steps and averaged two scans. Spectra were buffer-subtracted and converted to mean residue ellipticity (θ), and the secondary structure was estimated by K2D (47).

**NMR Spectroscopy**—15 ml of fresh LB broth containing the appropriate antibiotics was inoculated with 1:200 volume of starter culture and grown an additional ~8 h at 37 °C. The cells were then pelleted by centrifugation at 2000 rpm for 10 min and resuspended into 150 ml of M9 minimal medium (M9 salts, 0.2% glucose, 0.1% NH_{4}Cl, 2 mM MgSO_{4}, 100 μM CaCl_{2}, 10 μg/ml thiamine, 10 μg/ml biotin, and 10 μg/ml FeSO_{4}) supplemented with the same antibiotics. Following overnight growth at 37 °C, the cells were similarly pelleted and resuspended into 1.5 liter of M9 minimal medium containing 0.1% ^{15}NH_{4}Cl (Cambridge Isotope Laboratories, Inc.). After reaching a density of A_{600} ~1, induction was accomplished with 1 mM isopropyl-1-thio-β-D-galactopyranoside, and growth continued at 20 °C overnight. Proteins were purified by Ni-NTA and gel filtration. For the generation of untagged D4, half of the protein eluted from the Ni-NTA column was subjected to TEV protease digestion. For the final gel filtration step, the buffer was 50 mM sodium phosphate (pH 6.8), 30 mM NaCl, and 10% (v/v) glycerol. Proteins were used at ~250 μM concentrations in 10% D_{2}O. Spectra were recorded at 298 K on a Bruker Avance-III HD 600 spectrometer (Bruker Corp.) equipped with a 5-mm triple resonance TXI cryoprobe and z axis gradient using standard Bruker pulse sequences with WATERGATE for water suppression (48).

**In Vitro Processive DNA Synthesis**—All mutants were constructed in pcDNA3.2/V5-DEST (Invitrogen) for cell-free expression using rabbit reticulocytes as described previously (21). Plasmids were purified by phenol-chloroform extraction and used at 1 μg/50 μl of in vitro transcription/translation reaction (Promega). The translated proteins were then used to assess processive DNA synthesis by the rapid plate assay as described previously (21, 32).

**Prediction of Protein Disorder**—The D4 protein sequence was uploaded to the web interface of DisEMBL (41).

**Data Presentation and Analysis**—Data presentations and analyses (curve fitting) were performed with Prism 5 (GraphPad Software, Inc.) and UCSF Chimera (49).

**Author Contributions**—M. N. designed the study and conducted most of the experiments. M. N. and R. P. R. analyzed the results and wrote the paper. H. G. conducted the processive DNA synthesis experiments and contributed to the writing of the paper.
A Tripeptide of D4 is Essential for Processive DNA Synthesis

Acknowledgments—We thank Profs. G. H. Cohen and R. Eisenberg for use of the Biacore X100 and Dr. Chunhua Yuan for assistance with the NMR. This study made use of the Campus Chemical Instrument Center NMR facility at Ohio State University.

References

1. Ellison, V., and Stillman, B. (2001) Opening of the clamp: an intimate view of an ATP-driven biological machine. Cell 106, 655–660
2. Johnson, A., and O’Donnell, M. (2005) Cellular DNA replicases: components and dynamics at the replication fork. Annu. Rev. Biochem. 74, 283–315
3. Kong, X. P., Onrust, R., O’Donnell, M., and Kuriyan, J. (1992) Three-dimensional structure of the β subunit of E. coli DNA polymerase III holoenzyme: a sliding DNA clamp. Cell 69, 425–437
4. Jeruzalmi, D., O’Donnell, M., and Kuriyan, J. (2002) Clamp loaders and sliding clamps. Curr. Opin. Struct. Biol. 12, 217–224
5. Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell 79, 1233–1243
6. Kelman, Z. (1997) PCNA: structure, functions and interactions. Oncogene 14, 629–640
7. Gottlieb, J., Marcy, A. I., Coen, D. M., and Challberg, M. D. (1990) The herpes simplex virus type 1 UL42 gene product: a subunit of DNA polymerase that functions to increase processivity. J. Virol. 64, 5976–5987
8. Kiehl, A., and Dorsky, D. I. (1991) Cooperation of EBV DNA polymerase and EA-D(BMRF1) in vitro and colocalization in nuclei of infected cells. Virology 184, 330–340
9. Tsurumi, T., Daikoku, T., Kurachi, R., and Nishiyama, Y. (1993) Functional interaction between Epstein-Barr virus DNA polymerase catalytic subunit and its accessory subunit in vitro. J. Virol. 67, 7648–7653
10. Agulnick, A. D., Thompson, J. R., Iyengar, S., Pearson, G., Abassi, D., and Ricciardi, R. P. (1993) Identification of a DNA-binding protein of human herpesvirus 6, a putative DNA polymerase stimulatory factor. J. Gen. Virol. 74, 1003–1009
11. Weiland, K. L., Oien, N. L., Homa, F., and Wathen, M. W. (1994) Functional analysis of human cytomegalovirus polymerase accessory protein. Virus Res. 34, 191–206
12. Lin, K., and Ricciardi, R. P. (1998) The 41-kDa protein of human herpesvirus 6 specifically binds to viral DNA polymerase and greatly increases DNA synthesis. Virology 250, 210–219
13. Lin, K., Dai, C. Y., and Ricciardi, R. P. (1998) Cloning and functional analysis of Kaposis's sarcoma-associated herpesvirus DNA polymerase and its processivity factor. J. Virol. 72, 6228–6232
14. Moss, B. (2013) Poxvirus DNA Replication. Cold Spring Harbor Perspect. Biol. 5, 1–12
15. Upton, C., Stuart, D. T., and McFadden, G. (1993) Identification of a poxvirus gene encoding a uracil DNA glycosylase. Proc. Natl. Acad. Sci. U.S.A. 90, 4518–4522
16. Stuart, D. T., Upton, C., Higman, M. A., Niles, E. G., and McFadden, G. (1993) A poxvirus-encoded uracil DNA glycosylase is essential for virus viability. J. Virol. 67, 2503–2512
17. Mills, A. K., Carpenter, M. S., and DeLange, A. M. (1994) The vaccinia virus-encoded uracil DNA glycosylase has an essential role in viral DNA replication. Virology 198, 504–513
18. Ishii, K., and Moss, B. (2001) Role of vaccinia virus A20R protein in DNA replication: construction and characterization of temperature-sensitive mutants. J. Virol. 75, 1656–1663
19. Klemperer, N., McDonald, W., Boyle, K., Unger, B., and Traktman, P. (2001) The A20R protein is a stoichiometric component of the processive form of vaccinia virus DNA polymerase. J. Virol. 75, 12298–12307
20. Stanitsa, E. S., Arps, L., and Traktman, P. (2006) Vaccinia virus uracil DNA glycosylase interacts with the A20 protein to form a heterodimeric processivity factor for the viral DNA polymerase. J. Biol. Chem. 281, 3439–3451
21. Druck Shudoňsky, A. M., Silverman, J. E., Chattopadhyay, D., and Ricciardi, R. P. (2010) Vaccinia virus D4 mutants defective in processive DNA synthesis retain binding to A20 and DNA. J. Virol. 84, 12325–12335
22. Boyle, K. A., Stanitsa, E. S., Greseth, M. D., Lindgren, J. K., and Traktman, P. (2011) Evaluation of the role of the vaccinia virus uracil DNA glycosylase and A20 proteins as intrinsic components of the DNA polymerase holoenzyme. J. Biol. Chem. 286, 24702–24713
23. Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., and Tainer, J. A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. EMBO J. 17, 5214–5226
24. Schormann, N., Grigorian, A., Samal, A., Krishnan, R., DeLucas, L., and Chattopadhyay, D. (2007) Crystal structure of vaccinia virus uracil-DNA glycosylase reveals dimeric assembly. BMC Struct. Biol. 7, 45
25. Ishi, K., and Moss, B. (2002) Mapping interaction sites of the A20 protein component of the vaccinia virus DNA replication complex. Virology 303, 232–239
26. Sèbe, C., Gabel, F., Gutsche, I., Ivanov, I., Burmeister, W. P., Iseni, F., and Tarbouriech, N. (2013) Low-resolution structure of vaccinia virus DNA replication machinery. J. Virol. 87, 1679–1689
27. Obradovic, Z. (2002) Intrinsic disorder and protein function. Annu. Rev. Biochem. 71, 343–388
28. Nuth, M., Huang, L., Saw, Y. L., and Ricciardi, R. P. (2013) Design of potent poxvirus inhibitors of the heterodimeric processivity factor required for viral replication. J. Med. Chem. 56, 3235–3246
29. Bhuyan, A. K. (2002) Protein stabilization by urea and guanidine hydrochloride. Biochemistry 41, 13386–13394
30. Schormann, N., Banerjee, S., Ricciardi, R., and Chattopadhyay, D. (2015) Binding of undamaged double stranded DNA to vaccinia virus uracil-DNA glycosylase. BMC Struct. Biol. 15, 10
31. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. Nature 384, 87–92
32. Rachofsky, E. L., Seibert, E., Stivers, J. T., Osman, R., and Ross, J. B. (2001) Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine. Biochemistry 40, 957–967
33. Wright, P. E., and Dynson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J. Mol. Biol. 293, 321–331
34. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function. Biochemistry 41, 6573–6582
35. Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003) Protein disorder prediction: implications for structural proteomics. Structure 11, 1453–1459
42. Radhakrishnan, I., Pérez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator: coactivator interactions. Cell 91, 741–752
43. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. 6, 197–208
44. Kapust, R. B., Tözsér, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D., and Waugh, D. S. (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. Protein Eng. 14, 993–1000
45. Gasteiger, E. H. C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) in The Proteomics Protocols Handbook (Walker, J. M., ed.) pp. 571–607, Humana Press, Totowa, NJ
46. Myszka, D. G. (2000) Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE. Methods Enzymol. 323, 325–340
47. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. 32, W668–W673
48. Sklenar, V., Peterson, R. D., Rejante, M. R., Wang, E., and Feigon, J. (1993) 2-Dimensional triple-resonance HCNCH experiment for direct correlation of ribose H1′ and base H8, H6 protons in C-13, N-15 labeled RNA oligonucleotides. J. Am. Chem. Soc. 115, 12181–12182
49. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612