The vaccinia virus DNA polymerase structure provides insights into the mode of processivity factor binding

Nicolas Tarbouriech¹, Corinne Ducournau², Stephanie Hutin¹, Philippe J. Mas³, Petr Man⁴,⁵, Eric Forest¹, Darren J. Hart¹, Christophe N. Peyrefitte²,⁶, Wim P. Burmeister¹ & Frédéric Iseni²

Vaccinia virus (VACV), the prototype member of the Poxviridae, replicates in the cytoplasm of an infected cell. The catalytic subunit of the DNA polymerase E9 binds the heterodimeric processivity factor A20/D4 to form the functional polymerase holoenzyme. Here we present the crystal structure of full-length E9 at 2.7 Å resolution that permits identification of important poxvirus-specific structural insertions. One insertion in the palm domain interacts with C-terminal residues of A20 and thus serves as the processivity factor-binding site. This is in strong contrast to all other family B polymerases that bind their co-factors at the C terminus of the thumb domain. The VACV E9 structure also permits rationalization of polymerase inhibitor resistance mutations when compared with the closely related eukaryotic polymerase delta-DNA complex.
Poxviruses (members of the *Poxviridae* family) are large double-stranded DNA viruses that replicate exclusively in the cytoplasm of the infected cell. Viral DNA synthesis takes place in perinuclear sites called viral factories and depends on virus-encoded proteins. Poxviruses produce a number of non-essential enzymes involved in DNA precursor metabolism as well as essential proteins that are located at the replication fork. For vaccinia virus (VACV), the prototype and most-studied member of the *Poxviridae* family, the essential replication proteins include: E9, the catalytic subunit of the DNA polymerase; D4, a uracil-DNA glycosylase which together with A20 forms the heterodimeric processivity factor; D5, a hexameric nucleoside triphosphatase; and I3, a single-stranded DNA-binding protein. Other members of the virally encoded replication machinery include G5, a FEN-family endonuclease; A50, a DNA ligase and H5, an abundant hub protein.

For vaccinia virus (VACV), the prototype and most-studied member of the *Poxviridae* family, the essential replication proteins include: E9, the catalytic subunit of the DNA polymerase; D4, a uracil-DNA glycosylase which together with A20 forms the heterodimeric processivity factor; D5, a hexameric nucleoside triphosphatase; and I3, a single-stranded DNA-binding protein. Other members of the virally encoded replication machinery include G5, a FEN-family endonuclease; A50, a DNA ligase and H5, an abundant hub protein.

A low-resolution model of the VACV DNA polymerase holoenzyme E9/A20/D4 highlighted the elongated shape of the complex with a 150 Å separation between the DNA-binding sites of E9 and D4. A20 links both enzymes, and the DNA-binding properties of D4 are believed to increase the association of E9 with the genome template thus rendering the polymerase processive. Recently, high-resolution structures of the D4/A20 interface (D4/A20) and of D4/A20 bound to a 10-mer DNA duplex containing an abasic site resulting from the cleavage of an uracil base were obtained. These data further extend our knowledge on the processivity factor assembly and how DNA synthesis and base excision repair are coupled. However, structural information concerning the DNA polymerase and its interaction with A20 is still missing.

Over the years, a number of genetic and biochemical studies have characterized E9 (reviewed by Czernecky and Traktman). The enzyme is a member of the DNA polymerase family possessing DNA polymerase and 3’-5’ proofreading exonuclease activities. E9 alone was shown to be distributive under physiological conditions (adding fewer than 10 nucleotides per binding event) unless bound to its heterodimeric cofactor D4/A20.

### Table 1 Data collection and refinement statistics

|                         | E9-WT native | E9-WT Pb2⁺ | E9-WT Gd³⁺ | E9-WT Mn2⁺ | E9-exo⁻minus |
|-------------------------|--------------|------------|------------|------------|--------------|
| **Data collection**     |              |            |            |            |              |
| Beamline                | ESRF ID23-1  | ESRF BM14  | ESRF BM14  | ESRF ID23-1| ESRF ID23-1  |
| Space group             | P₂₁,₂₁      | P₂₁,₂₁     | P₂₁,₂₁     | P₂₁,₂₁     | P₂₁,₂₁       |
| Cell parameters (Å)     | 133.4 133.4 230.5 | 133.7 133.7 229.6 | 133.6 133.6 229.8 | 134.0 134.0 230.2 | 133.5 133.5 229.5 |
| Wavelength (Å)          | 1.0714       | 1.2724     | 1.4226     | 1.0971     | 1.2724       |
| Resolution range (Å)    | 1.040-1.367 | 1.040-1.367 | 1.040-1.367 | 1.040-1.367 | 1.040-1.367 |
| No. of observed reflections | 332899 (22308) | 176231 (35375) | 236574 (24674) | 199915 (19915) | 188465 (26270) |
| No. of unique reflections | 63014 (4369) | 23925 (4779) | 43418 (4429) | 59597 (5772) | 57789 (8176) |
| Completeness (%)        | 99.9 (99.9)  | 99.4 (98.1)| 99.5 (98.9)| 99.1 (99.2)| 99.0 (97.1)  |
| Multiplicity            | 0.5 (1.1)    | 7.4 (7.4)  | 5.4 (5.6)  | 3.4 (3.5)  | 3.3 (3.2)    |
| Mean (I/σ(I))           | 6.9 (65.3)   | 13.1 (44.0)| 10.4 (40.5)| 9.1 (81.0) | 7.4 (65.3)   |
| Rmerge (%)              | 0.999 (0.862)| 0.997 (0.944)| 0.996 (0.942)| 0.996 (0.679)| 0.996 (0.631)|
| Mosaicity (°)           | 0.07         | 0.12       | 0.06       | 0.03       | 0.06         |
| Overall B factor (Wilson plot) (Å²) | 49.0   | 66.1      | 45.6      | 58.8      | 43.8         |

**Model refinement and composition**

|                         | E9-WT native | E9-WT Pb2⁺ | E9-WT Gd³⁺ | E9-WT Mn2⁺ | E9-exo⁻minus |
|-------------------------|--------------|------------|------------|------------|--------------|
| No. of reflections, working set | 59939       | 56686      | 54984      | 54984      | 54984        |
| No. of reflections, test set | 3075        | 2911       | 2805       | 2805       | 2805         |
| Final Rcryst (%)        | 0.186        | 0.183      | 0.185      | 0.185      | 0.185        |
| Final Repro (%)         | 0.186        | 0.183      | 0.185      | 0.185      | 0.185        |
| No. of non-H atoms      | 59939        | 56686      | 54984      | 54984      | 54984        |
| Protein (residues)      | 8188         | 8180       | 8173       | 8173       | 8173         |
| Ligand                  | 71           | 75         | 71         | 71         | 71           |
| Water                   | 219          | 215        | 214        | 214        | 214          |
| Total                   | 8478         | 8470       | 8458       | 8458       | 8458         |
| Model composition       |              |            |            |            |              |
| Protein (residues)      | 999          | 999        | 999        | 999        | 999          |
| MES                     | 3            | 3          | 3          | 3          | 3            |
| HEPES                   | 1            | 1          | 1          | 1          | 1            |
| DTT                     | 1            | 1          | 1          | 1          | 1            |
| Glycerol                | 2            | 2          | 2          | 2          | 2            |
| Mn²⁺                    | 0            | 4          | 0          | 0          | 0            |
| RMS deviations          |              |            |            |            |              |
| Bond lengths (Å)        | 0.015        | 0.011      | 0.011      | 0.011      | 0.011        |
| Bond angles (°)         | 1.853        | 1.553      | 1.617      | 1.617      | 1.617        |
| Average B factor (Å²)   | 74           | 72         | 78         | 78         | 78           |
| Ramachandran plot (%)   |              |            |            |            |              |
| Favored                 | 92.4         | 94.3       | 93.1       | 93.1       | 93.1         |
| Additionally allowed    | 6.3          | 4.5        | 5.8        | 5.8        | 5.8          |
| Outliers                | 1.3          | 1.2        | 1.1        | 1.1        | 1.1          |

*Values in parentheses correspond to the highest-resolution shell

*Estimated R_repro = \( R_{repro} = \sqrt{(N(N-1))}/2 \), where N is the data multiplicity
The DNA polymerase was also shown to catalyze annealing of single-stranded DNA, an activity not found in other DNA polymerase family B members. The end-joining reaction requires the 3′–5′ exonuclease activity of E9 that degrades the extremities of dsDNA to create 5′-ssDNA overhangs. Sequence alignments with other DNA polymerases identified E9-specific insertions but have not yet been correlated with precise functions.

E9 is the target of several inhibitors such as aphidicolin (aph), phosphonoacetic acid (PAA), cytosine arabinoside (AraC), and CMX001, a cidofovir (CDV) derivative, which is in advanced development for the treatment of smallpox. These compounds have been used to select and characterize resistance mutations located in the VACV polymerase.

Here we present the 2.7 Å crystal structure of the full-length VACV DNA polymerase and the characterization of the interface between E9 and its processivity factor subunit A20. The data allow us to explore the role of E9-specific inserts and to position E9 in a global model of the DNA polymerase holoenzyme which differs from other family B polymerases.

**Results**

**Crystal structure of E9 the VACV DNA polymerase.** E9 crystallized in space group P3_121 with one molecule in the asymmetric unit. The structure was solved at 2.7 Å resolution using the MIRAS method (Table 1). We observed the classical palm, thumb, finger, exonuclease, and N-terminal domains of a family B polymerase in an open conformation (Fig. 1a). Only 10 out of the 1006 residues could not be modeled. A previous sequence analysis of E9 allowed the delimitation of “poxvirus-specific” inserts that we can now redefine based on flexible structural alignments with other family B polymerases. Three poxvirus-specific inserts are clustered on one side of the molecule, corresponding to insert 0 (aa 67–82), insert 3 (aa 572–610), and insert 4 (aa 708–743) (Fig. 1b), whereas insert 1 (aa 208–233) is located on the opposite side (Fig. 1c). Insert 2 (aa 356–432) is located on the back of E9 in the exonuclease domain (Fig. 1c) and forms a little 6-stranded β-barrel with a greek key fold. The previously assigned insert 5 does not exist whilst a new insert in the N-terminal domain is identified and denominated as insert 0 (Fig. 1b).

The crystal packing involves mainly the periphery of the molecule, implicating N-terminal, exonuclease, palm, and thumb domains. The helix of insert 3 forms an important contact as it is bound in a hydrophobic cavity of a symmetry-related molecule mainly comprising residues of the exonuclease domain (i.e., residues 190–200 and 227–233). The temperature factors indicate a high mobility of insert 2 (which appears to be in loose contact with the body of the polymerase) and of the thumb domain. A high mobility of the thumb domain is generally observed for family B polymerase. Still, for E9 electron density is visible for all the C-terminal residues.

A calculated electrostatic potential of E9 shows extensive positively charged surfaces on both faces of the molecule (Fig. 1d). On the front face (defined by the inferred location of the polymerase active site), these coincide with areas of conserved sequence within orthopoxviruses (Fig. 1e). On the back, however, little sequence conservation is observed.

**Identification of a soluble A20 domain interacting with E9.** Within the VACV DNA polymerase holoenzyme, A20 forms a link between D4 and the catalytic subunit of the polymerase E9. While the first 50 residues of A20 were shown to interact with D4, the low-resolution structure of E9/A20/D4 suggests that the A20 C-terminal region may be involved in E9 interaction. We used the ESPRIT technology to identify soluble purifiable C-terminal fragments of A20. A random library of 5′ truncations of the full-length A20R gene was generated using exonuclease III and mung bean nuclease. Approximately 4600
truncation mutants were isolated by robotic colony picking and tested for expression of soluble protein in an *E. coli* colony-based screen on nitrocellulose filters. Thirty clones exhibited levels of purified protein visible by Coomassie blue-stained SDS-PAGE. Constructs yielding highly soluble A20 fragments were sequenced and revealed nine unique clones expressing the last 116–148 residues of A20 (Supplementary Fig. 1), of which the clone containing the last 123 residues of the protein (A20 C-ter) was selected for further experiments.

To obtain structural information on this domain, the A20 C-ter fragment was analyzed by SAXS (Supplementary Table 2). The pronounced maximum of the Kratky plot indicated that the construct forms essentially a compact folded structure (Supplementary Fig. 2d) and mutant protein constructs were similar, hydrophobic residues as diamonds, charged residues as triangles and pentagons. The alignment of the last amino acids of A20 using representative viruses from each genus of the Chordopoxvirinae subfamily: VACV, CPXV (cowpox virus), MPXV (monkeypox virus), CMLV (camelpox virus), ECTV (ectromelia virus), VARV (variola virus), Y KPV (yokapox virus), LSDV (lumpy skin disease virus), TANV (tanapox virus), SWPV (swinepox virus), MYXV (myxoma virus), MO CV (molluscum contagiosum virus), DPV (deerpox virus), ORFV (ORF virus), CNPV (canarypox virus). Conserved residues are in red. The predicted secondary structure using MLRC and its reliability on a 0–9 scale are shown at the bottom of the alignment (ACC). Secondary structure prediction suggested that the last 27 residues of A20 C-ter, together with a twofold molar excess of A20 C-ter, a first peak is eluted at 12.9 mL and a second one at 17.5 mL (Fig. 2a). SDS-PAGE analysis of the eluted protein fractions showed that the second peak contains the expected excess of free A20 C-ter (Fig. 2b), whereas part of the A20 C-ter is co-eluted with E9 in the first peak indicating an interaction between both proteins.

To confirm the interaction between E9 and A20 C-ter, surface plasmon resonance (SPR) was performed in which A20 C-ter (biotinylated in vivo via its BAP tag) was immobilized on streptavidin-coated chips. Serial dilutions of E9 were injected and analysis of resulting sensorgrams yielded a $K_D$ of 23 nM for the interaction (Supplementary Fig. 3a).

**Identification of residues involved in A20/E9 interaction.** As it was not possible to crystallize the complex between E9 and A20 C-ter, we aimed to characterize the interface using biophysical techniques. The stability of the E9/A20/D4 holoenzyme under high salt conditions (NaCl > 750 mM) observed by us and others suggested that the E9/A20 interaction may be significantly hydrophobic in nature. The alignment of sequence homologs to the A20 C-ter fragment indicated an $\alpha$-helix content of about 45% (Supplementary Fig. 2d) and mutant protein constructs were similar, hydrophobic residues as diamonds, charged residues as triangles and pentagons. A point mutation in A20 C-ter abrogates binding to E9. A20 C-ter-Phe414Ala was incubated with WT E9 before injection onto SEC. The eluted fractions were analyzed as in b.

![Image](https://example.com/image.png)
affected complex formation between E9 and A20 C-ter. As shown in Fig. 2e, E9 and A20 C-ter-Phe414Ala did not co-purify showing that the interaction was lost. This was confirmed by SPR experiments where the residual interaction between E9 and A20 C-ter-Phe414Ala was too weak to determine a $K_D$ (Supplementary Fig. 3b).

The Phe414Ala mutation was then introduced into the full-length VACV A20. WT and mutant proteins were expressed in insect cells and purified in complex with D4 as previously described. The ability of both heterodimeric complexes to bind to WT E9 was assessed by SPR. WT D4/A20 interacts with E9 with a $K_D$ of 8 nM (Supplementary Fig. 3c) in agreement with previous results. In contrast, D4/A20 Phe414Ala showed much weaker binding and a marked bi-exponential dissociation phase so that a $K_D$ could not be calculated (Supplementary Fig. 3d). Altogether, these data indicated that the putative $\alpha$-helix present

Fig. 3 Peptides protected by the E9 A20 C-ter interaction determined by H/D exchange MS. a Peptides from E9, identified by LC-MS/MS, are represented as bars above the primary sequence. The secondary structure, derived from the crystal structure, is shown underneath the sequence (in gray residues not seen in the structure). The level of protection of individual peptides, as determined by comparing the % D incorporation for free E9 with the one for E9 bound to A20 C-ter, is color coded according to the scale bar. Highly protected areas are shown in red, whereas peptides becoming more exposed upon complex formation are shown in blue. The highly protected region involves the $\alpha$-helix of insert 3 (residues 577–590). b H/D exchange results are mapped onto the crystal structure of E9. Protected residues are color coded as in a.
at the C-terminal extremity of A20 is involved in the interaction with E9.

To identify the interaction surface of A20 C-ter on E9, we used hydrogen–deuterium exchange mass spectrometry under native conditions. We analyzed the E9/A20 C-ter complex by measuring deuterium exchange for 107 partly overlapping peptides from E9, alone, or in complex with A20 C-ter. These correspond to 75% of the E9 primary sequence (Fig. 3a). The comparison of the level of deuteration highlighted one distinct region that showed a strong reduction in deuterium exchange in the complex, indicating protection upon complex formation. Interestingly, all the corresponding peptides cover the α-helix of insert 3 (Figs. 1b and 3b). We also observed peptides displaying increasing deuterium exchange upon A20 C-ter binding, which cluster in the exonuclease domain of E9 and could indicate conformational changes due to the interaction with A20 C-ter (Fig. 3a, b).

In order to confirm the involvement of E9 insert 3 in A20 binding, three mutants of conserved residues in the α-helix were produced. Hydrophobic residues (Leu578 and Ile582) were mutated to alanine (mutant E9-578-582), charged residues (Glu580 and Glu581) were mutated to oppositely charged arginine (E9-580-581) and Gln585-Gln589 and Leu586-Leu588 residues were mutated to Ala and Ser, respectively (E9-585-6-8-9) (Fig. 4a, c, e). When E9-578-582 was incubated with WT A20 C-ter and loaded on SEC, both proteins eluted separately (Fig. 4b), indicating that the interaction is largely reduced; this was subsequently confirmed by SPR (KD = 390 nM, i.e., 17-fold reduction, Supplementary Fig. 3e). When the charged residues were mutated (E9-580-581), the interaction with A20C-ter was still observed on SEC, likewise for the quadruple mutant E9-585-6-8-9 (Fig. 4d, f). However, SPR experiments showed reduced binding; a 10-fold decrease in KD for the E9-580-581 mutant and fourfold decrease for the E9-585-6-8-9 mutant compared to WT E9 (Supplementary Fig. 3f, g). Thus, we conclude that the hydrophobic residues (Leu578 and Ile582) on the N-terminal side of the insert 3 α-helix are key contacts in E9/A20 complex formation, with neighboring residues involved to a lesser extent, although still enough to permit co-purification of the complex on SEC.

**Related polymerase structures and modeling of E9/DNA complex.** In order to find the most closely related structures to E9, a PDB search using a block-wise structural alignment was used to overcome the inherent flexibility of polymerases. This yielded family B polymerases including E. coli DNA polymerase II (PDB 3k57), archaeal family B polymerases (PDB 2xh4), yeast polymerase δ (pol δ, PDB 3iap), herpes simplex virus (HSV) polymerase (PDB 2gv9), and eukaryotic DNA polymerase α (pol α, for example PDB 4q5v) as the most similar ones. Bacteriophage RB69 polymerase and eukaryotic polymerase ε are more distantly related.33,35,36 As the thumb domain movements contribute most to the conformational variability of family B polymerases, a structural alignment of E9 excluding the thumb domain was carried out, which identified yeast pol δ as the most closely related protein (Supplementary Table 1). The close structural relationship between E9 and yeast pol δ was further confirmed when separate E9 domains were superposed individually (Supplementary Table 1).

The published family B polymerase structures in complex with DNA oligonucleotides, with or without an incoming nucleotide mimicking elongation or editing modes, indicate that considerable domain movements occur upon DNA binding leading to a closure of the structure compared to the apo forms. The yeast pol δ structure with bound template and complementary DNA (PDB 3iap) was used to model E9 in elongation mode. We performed SAXS experiments (Supplementary Table 2) using an isomorphous exominus mutant of E9 (Table 1 and Supplementary Fig. 4a, b) bound to a 29-mer DNA hairpin to check if the best model would be obtained by adjusting only the position of the thumb domain, or if a movement of each individual domain would be required. Complex formation was confirmed by the decrease of the radius of gyration from 3.83 to 3.45 nm upon DNA binding (Supplementary Table 2). The comparison of the SAXS curve of the E9exo/minus/DNA complex with the calculated scattering of different E9 models (Supplementary Fig. 5a–c) showed that when individual domains of E9 are adjusted (Supplementary Fig. 5c), the theoretical curve fits better (and the χ² is lower) than when only the thumb domain is adjusted (Supplementary Fig. 5b) supporting the generalized domain movements upon DNA binding.

Likewise, the enzyme in proofreading or editing mode can be modeled based on the structure of an archaeal polymerase in editing mode37 (PDB 2xh4, Supplementary Fig. 4c). Larger differences in domain orientation between E9 and the archaeal enzyme make the model globally less reliable, especially since SAXS data are not available. However, the modeled DNA fits very precisely into the exonuclease active site of E9 (Supplementary Fig. 4d), in particular regarding the position of the catalytic Mg²⁺.
ions replaced by Mn$^{2+}$ in one of the E9 structures (Table 1 and Supplementary Fig. 4e).

**Discussion**

The high-resolution structure of the catalytic subunit of the VACV DNA polymerase allows comparison of E9 with other replicative DNA polymerases. Using different criteria such as the global structural similarity of the polymerase domain and structural superpositions of individual subdomains, we found that E9 appears to have evolved for cofactor binding, whereas the poxvirus-specific inserts 0 and 4 appear to buttress insert 3 (Fig. 1b).

Insert 2 forms a small 6-stranded $\beta$-barrel domain that is also found in ATP synthase F1, EF-Tu, Gar1, and other proteins$^{38,39}$. It only shows weak sequence conservation between poxviruses (Fig. 1c) and the absence of strong electrostatic features (Fig. 1d) make it unlikely that this domain interacts with nucleotides. Interestingly, insert 2 carries a number of resistance mutations together with mutation Ala684Val. PAA$^{hs}$, aph$^{r}$, PAA$^{hs}$, araC$^{r}$, araA$^{hs}$ Red Interference with the rotation of a base in the template backbone next to aph-binding site 68

Table 2 E9 drug-resistant and temperature-sensitive mutants

| Mutation | Virus | Effect | Color code | Proposed mechanism | Ref. |
|----------|-------|--------|------------|-------------------|------|
| Phe171Ser | VACV | araC$^{r}$, araA$^{r}$, aph$^{hs}$ (when combined with mutant Cys356Tyr or Gly372Asp or Gly380Ser) | Yellow | Modification of 3$'$-nucleotide binding at exonuclease site | 21 |
| $\Delta$Lys174 | VACV | CDV$^{r}$ | Yellow | Modification of 3$'$-nucleotide binding at exonuclease site | 20 |
| His185Tyr | VACV | ts Dts83 | Black | Destabilization of the exonuclease domain | 67 |
| Ala314Thr/Val | VACV, CMLV$^{a}$, MPV$^{b}$ | CDV$^{r}$ and cross-resistance to other nucleoside phosphonate drugs. Stronger drug-resistance together with mutation Ala684Val. PAA$^{hs}$ | Orange | Mutation in beta-hairpin; may modulate guidance from elongation to editing mode | 20,22-26 |
| Ser338Phe | VACV | CDV$^{r}$ | Orange | Indirect effect on complementary strand binding or strand switching between elongation and editing mode? | 25 |
| Cys356Tyr | VACV | PAA$^{r}$, aph$^{hs}$ | Green | Effect on insert 2 $\beta$-finger interaction? | 21 |
| Gly372Asp | VACV | PAA$^{r}$, aph$^{hs}$ | Green | Effect on insert 2 $\beta$-finger interaction? | 21 |
| Gly380Ser | VACV | PAA$^{r}$, aph$^{hs}$ | Green | Effect on insert 2 $\beta$-finger interaction? | 21,68 |
| Gly392Asp | VACV | ts NG26 | Black | Destabilization of insert 2 | 27 |
| Ala498Thr/Val | VACV | ap$^{h}$, PAA$^{hs}$, araC$^{hs}$, araA$^{hs}$ | Red | Interference with the rotation of a base in the template required for aph binding | 68 |
| Glu611Lys | VACV | ts Cts42 | Black | Destabilization of insert 3 or of the elongation site | 44 |
| Leu670Met | VACV | aph$^{r}$ | Red | Indirect effect on the binding of the template backbone next to aph-binding site | 22-24,26 |
| Ala684Val/Thr | VACV, CMLV$^{a}$, MPV$^{b}$ | CDV$^{r}$ and cross-resistance to other nucleoside phosphonate drugs. Stronger drug-resistance together with mutation Ala314Thr | Red | Indirect effect on the binding of the template backbone in the elongation site | 22-24,26 |
| Ser686Asn | VACV | ts, Dts20 | Black | Destabilization of insert 3 or of the elongation site | 67 |
| Thr688Ala | VACV | CDV$^{r}$ and cross-resistance to other nucleoside phosphonate drugs when associated with Ala314Thr mutation. PAA$^{hs}$, araC$^{hs}$ | Red | Indirect effect on template backbone binding in the elongation site | 22 |
| Thr311Le | VACV | CDV$^{r}$ and cross-resistance to other nucleoside phosphonate drugs | Pink | Modulation of domain movements by a modification at the thumb-palm domain connection and a changed interaction with the complementary strand | 24 |
| Ser851Tyr | VACV | CDV$^{r}$ and cross-resistance to other nucleoside phosphonate drugs when combined with Ala684Val mutation | Pink | Modulation of domain movements by a modification of the thumb-palm domain interface | 23 |

$c$, drug-resistant; $hs$, drug-hypersensitivity; $ts$, temperature-sensitive

$^{a}$Camepox virus

$^{b}$Monkeypox virus

Insert 2 forms a small 6-stranded $\beta$-barrel domain that is also found in ATP synthase F1, EF-Tu, Gar1, and other proteins$^{38,39}$. It only shows weak sequence conservation between poxviruses (Fig. 1c) and the absence of strong electrostatic features (Fig. 1d) make it unlikely that this domain interacts with nucleotides. Interestingly, insert 2 carries a number of resistance mutations together with mutation Ala684Val. PAA$^{hs}$, aph$^{r}$, PAA$^{hs}$, araC$^{r}$, araA$^{hs}$ Red Interference with the rotation of a base in the template backbone next to aph-binding site 68

Analysis of the E9 structure does not explain the unique role of the enzyme in recombination$^{17}$, as no obvious domain could be related to such an activity. Mechanistically, the reaction was shown to require the 3$'$-to-5$'$ exonuclease activity of the polymerase$^{18}$. It is intriguing that insert 2 is located within the essential exonuclease domain. However, poxvirus DNA synthesis and recombination appear to be tightly linked processes and, to date, genetic and biochemical studies have failed to isolate E9 mutants with functionally separate activities.

A low-resolution structure of the holoenzyme obtained by SAXS is available$^{11}$. The E9 structure can be positioned unambiguously in the large part of this envelope but the orientation of the disk-shaped E9 around its short axis remains unclear (Fig. 5a). Using manual fitting based on the position of the A20-binding site on E9 (i.e., the $\alpha$-helix from insert 3), the orientation of E9 can be obtained with an accuracy of about $\pm40^\circ$. Using the
In general, family B polymerases require co-factors for processivity. Figure 5b-e illustrates the evolution of various co-factors bound to their cognate polymerase. For the bacteriophage RB69 and archaeal polymerases, a direct interaction between the tri-meric sliding clamp processivity factor and the extreme C-terminus of the DNA polymerase is observed (Fig. 5c). Similarly, HSV polymerase interacts with its PCNA-related processivity factor (UL42) through the C-terminal peptide of the polymerase (Fig. 5d). Eukaryotic polymerase δ (pol3 in yeast) carries an elongated C-terminal domain that binds to PCNA (Fig. 5e). To date, it is not possible to infer any relationship between VACV A20 and polymerase δ co-factor subunits or any other protein at the level of sequence or secondary structure. Understanding the evolution of A20 as a necessary subunit of the E9 holoenzyme will require the high-resolution structure of the full-length protein.

There are two possibilities for the orientation of the VACV holoenzyme at the replication fork: either D4 interacts with the newly synthesized dsDNA (Fig. 5f) or D4 binds to the parental ss/dsDNA (Fig. 5g, h). In the latter case, as there is no information on the actual length of the DNA between the DNA-binding site of D4 and the active site of E9, it is possible that the D5 helicase–primase is located between both binding sites (Fig. 5g). It cannot be excluded that D5 moves ahead of the polymerase holoenzyme with D4 binding to ssDNA (Fig. 5h) as D4 appears to bind ssDNA and dsDNA equally (competing with dCTP) are inhibitors blocking polymerase action, whereas AraC and cidofovir (both nucleoside analogs) terminate chain elongation after their integration into newly synthesized DNA. The modeled E9 structure in a closed, DNA-terminating conformation suggests three main resistance mechanisms: (i) changes in the 3′–nucleotide of the DNA (Fig. 6a and Supplementary Fig. 4d). (ii) Perturbation of template backbone binding in the exonuclease site facilitating the removal of chain terminators, and (iii) mutations affecting the switch from elongation to editing mode.

(i) The AraC resistance mutation Phe171Ser21, located in the exonuclease active site, targets a residue predicted to contact the 3′-nucleotide of the DNA (Fig. 6a and Supplementary Fig. 4d).
The prominent β-finger located in the exonuclease domain (residues 299–319) is a structural element, which seems important for switching between elongation and editing modes as it is able to contact the template strand during elongation and intervene in strand separation required for proofreading16. In E9, this domain harbors the principal CDV resistance mutation Ala314Thr/Val322–26 that facilitates excision of nucleoside analogs18. Indirectly, the Ser338Phe mutation23 may also affect the switch from elongation to editing mode. Likewise, Thr831Ile and Ser851Tyr mutations23,24 located at the interfaces of the thumb domain may influence the domain movements required for transition from elongation to editing mode facilitating indirectly the excision of nucleotide analogs. The Thr831Ile mutation might also interact directly with the complementary strand.

The resistance mutations raised against PAA21 are a class on their own (Fig. 6b). They are all located in the poxvirus-specific insert 2 domain and could only be explained by a movement of the finger domain upon binding of an incoming nucleotide as described for other family B polymerases17,18. This would bring the tip of the finger and insert 2 into contact (Fig. 6b). The same movement has been observed upon binding of the PAA-related molecule phosphonoformate to a modified RB69 polymerase mimicking the finger domain from HSV, which blocks the polymerase in a closed conformation19. Consequently, mutations Cys356Tyr and Gly380Ser that likely destabilize the hydrophobic core of the insert 2 domain, or Gly372Asp that may affect the position of the surface residue Trp411, could have an indirect effect on the interaction of insert 2 with the finger, which in turn might influence PAA binding in the active site.

Altogether, the high-resolution structure of the poxvirus DNA polymerase not only identifies the mode of processivity factor binding, but also permits an understanding of antiviral resistance mechanisms, although the dynamic character of the polymerase still requires further structural work in order to obtain more precise snapshots of the different functional states. The identification of the E9/A20 interface will facilitate the design of compounds or peptides that disrupt this interaction, whereas the high-resolution structure of the polymerase active site will accelerate the development of new antiviral drugs.

Methods

**E9 expression and purification.** VACV E9 (Copenhagen strain) was expressed in high five insect cells (Thermo Fisher Scientific) infected with a recombinant baculovirus carrying the E9L gene fused to a N-terminal His-tag and a TEV cleavage site11. Cell suspensions were grown in Express Five-SFM medium (Gibco) at 27 °C following protocols described in Trowitzsch et al.30. Cell pellets were resuspended in 10 volumes of equilibration buffer (50 mM Tris-HCl pH 7, 300 mM NaCl, 10 mM β-mercaptoethanol, and 10 mM imidazole) with complete, EDTA-free protease inhibitor cocktail (Roche). Cells were disrupted mechanically using a Potter-Elvehjem homogenizer. The lysate was clarified by centrifugation at 48,000 x g for 20 min at 4 °C. The supernatant was loaded onto a 5 mL HisTrap FF crude column (GE Healthcare) previously equilibrated with equilibration buffer. The column was washed with equilibration buffer containing 20 mM imidazole and proteins were eluted with a linear 20–250 mM imidazole gradient.

Fractions containing E9 were pooled and concentrated using Amicon centrifugal filter units (Millipore). The buffer was exchanged to equilibration buffer using an Econo-10 DG desalting column (Bio-Rad). The protein was digested overnight at 20 °C with papain and then loaded onto a 5 mL HisTrap FF crude column. E9 was recovered in the flow-through fraction and was concentrated prior to injection onto a Superdex 200 GL 10/300 column (GE Healthcare) equilibrated with gel filtration buffer (20 mM Tris-HCl pH 7, 300 mM NaCl, 4 mM dithiothreitol (DTT)). Fractions were analyzed by SDS-PAGE and stained with InstantBlue (Expedeon). The E9 mutants subsequently called exo18–18 (Asp166Ala + Glu168Ala) (exonuclease deficient by mutation of Mg2+ coordinating residues), E9–278–582 (Leu578Ala + Ile582Ala: aliphatic residues replaced by small residues) were produced by PCR-based site-directed mutagenesis (Supplementary Table 3) and expressed and purified as described above.

**Methods**

**E9 analysis of drug-resistance mutations.** Analysis of the apo structure of human pol α (PDB 5iud), in orange E9 after superposition of the palm and exonuclease domains onto the binary complex of human pol α, in blue the ternary complex of yeast pol δ, in purple E9 after superposition of the palm and exonuclease domains onto the binary complex of human pol α (residues 299–319) is a structural element, which seems important for switching between elongation and proofreading modes as it is able to contact the template strand during elongation and intervene in strand separation required for proofreading. In E9, this domain harbors the principal CDV resistance mutation Ala314Thr/Val322–26 that facilitates excision of nucleoside analogs. Indirectly, the Ser338Phe mutation may also affect the switch from elongation to editing mode. Likewise, Thr831Ile and Ser851Tyr mutations located at the interfaces of the thumb domain may influence the domain movements required for transition from elongation to editing mode facilitating indirectly the excision of nucleotide analogs. The Thr831Ile mutation might also interact directly with the complementary strand.

**Methods**

**Fig. 6 Analysis of drug-resistance mutations on E9.** a Locations of different mutants listed in Table 2 on the model of E9 in complex with DNA in elongation mode. Domains of the polymerase are color coded as in Fig. 1a. b Analysis of the finger domain movements: in magenta, the DNA-bound structure of human pol α (PDB 5iud), in orange E9 after superposition of the palm and exonuclease domains onto the binary complex of human pol α, in blue the ternary complex of yeast pol δ, in purple E9 after superposition of the palm and exonuclease domains onto the binary complex of human pol α (residues 299–319) is a structural element, which seems important for switching between elongation and editing modes as it is able to contact the template strand during elongation and intervene in strand separation required for proofreading. In E9, this domain harbors the principal CDV resistance mutation Ala314Thr/Val322–26 that facilitates excision of nucleoside analogs. Indirectly, the Ser338Phe mutation may also affect the switch from elongation to editing mode. Likewise, Thr831Ile and Ser851Tyr mutations located at the interfaces of the thumb domain may influence the domain movements required for transition from elongation to editing mode facilitating indirectly the excision of nucleotide analogs. The Thr831Ile mutation might also interact directly with the complementary strand.
E9 crystallization and data collection. The E9 protein was concentrated to 7 mg mL⁻¹ in 20 mM Tris-Cl pH 7, 300 mM NaCl, 4 mM DTT using Amicon 50 kDa Ultra Centrifugal filters. Initial crystallization conditions were identified in an initial screen. Crystals were obtained from conditions that were found to be in the Morphous screen (Molecular Dimensions) (10% PEG 4000, 20% glycerol, 100 mM MES-imidazole pH 6.5, 15 mM MgCl₂, 15 mM CaCl₂) using the EMBL Grenoble high-throughput crystallization facility and were refined manually to 9–11% PEG 3000, 20–25% glycerol, 100 mM MES-NAOH pH 6.25. Heavy atom derivates were obtained by soaking crystals for 2–24 h in reservoir solution supplemented with 1 mM of either Pb(CH₃COO)₂ or CdCl₂. Needle-like crystals were flash frozen in liquid nitrogen before data collection on ESRF beamlines ID31 for native data sets and BM14 for heavy atom derivatives. Helical data collection at medium resolution was transferred in their current form to overcome radiation damage. Mn⁺² derivatised crystals were obtained by soaking native crystals in the reservoir solution with an additional 5 mM MnCl₂.

Structure determination. The structure of E9-WT was determined by the SIRAS method using anomalous scattering. DXTS was used for data integration, AIMLESS was used for data reduction, SOLVE and RESOLVE were used for phase determination and improvement. A first model built with Buccaneer was further refined using cycles of manual inspection and building using COOT and restrained refinement with individual B factors but without translation/libration/screw refinement using REFMAC5. E9exo₅⁶₉ and Mn⁺² structures were isomorphous.

Structure analysis and modeling. The final E9 structure was compared to other DNA polymerases using the flexible structure alignment algorithm implemented in FATCAT. The insert 2 domain of E9 was modelled with CATH identified folds of the closed and DNA-bound conformations used the “super” structural alignment function of PyMOL (The PyMOL Molecular Graphics System, Version 1.4.1 Schrödinger, LLC). Figures were generated with PyMOL.

Identification of the A20 C-ter soluble construct. The full-length 1281 bp VACV A20R gene was subcloned into pESPRIT002, a pET9a-derivative encoding N-terminal His-tag and C-terminal BAP. The outlet cloning cassette was then cloned in-frame with the BAP-encoding sequence. The final E9 structure was compared to other DNA polymerases using the flexible structure alignment algorithm implemented in FATCAT. The insert 2 domain of E9 was modelled with CATH identified folds of the closed and DNA-bound conformations used the “super” structural alignment function of PyMOL (The PyMOL Molecular Graphics System, Version 1.4.1 Schrödinger, LLC). Figures were generated with PyMOL.

D4/A20 expression and purification. VACV D4 (fused to a N-ter His-tag and a TEV cleavage site) and full-length WT or mutant A20 were co-expressed in insect cells infected with a recombinant baculovirus as described in Sèze et al.²². Protein expression and purification were essentially performed following the protocols described for E9, except the elution step from the first nickel column, which used 100 mM imidazole. The final gel filtration step was performed on a Superdex 200 10/300 column (GE Healthcare) equilibrated in 25 mM Tris-Cl pH 7.5, 300 mM NaCl. The D4/A20 Phe414Ala mutant was generated by PCR-based site-directed mutagenesis (Supplementary Table 3).

Surface plasmon resonance. SPR acquisitions were carried out on a CMS sensorchips on a BIACore 3000 instrument (GE Healthcare). All experiments were performed in buffer containing 100 mM NaCl, 25 mM Tris-Cl pH 7.5 at a flow rate of 15 μL min⁻¹. For experiments involving A20 C-ter, about 2500 resonance units (RU) of streptavidin (Sigma-Aldrich) were immobilized on the EDC-NHS activated surfaces. A20 C-ter (or A20 C-ter-Phe414Ala) was injected at 10 μg mL⁻¹ into one flow cell until ~1000 RU were reached. A second flow cell (without bound A20 C-ter) was used for background subtraction. Twofold serial dilutions (160–5 μM) of E9 in running buffer were injected during 180s (association phase) followed by a 150 s dissociation phase. Similarly, for experiments involving full-length D4/A20 complex or Phe414Ala mutant complex, E9-WT was immobilized and a twofold serial dilution (160–5 μM) of complex in running buffer was injected. Background subtracted signals were exported from the Biologic software (GE Healthcare) and imported into the LiberoOffice Calc (LiberoOffice.org) for curve fitting using the Solver function and figure preparation.

Small-angle X-ray scattering experiments. An E9exo₅⁶₉/DNA complex has been prepared by mixing E9exo₅⁶₉ with a 20% molar excess of a DNA 29-mer
forming an hairpin structure with five bases overhang at the 5’ end, 10 base-pairs and a 4 nucleotide loop (5’-AAGGCGGCTTGCTTTTCCACGACAAGC-3’) similar to the one used by Killilea and co-workers37. Fifty and a 4 nucleotide loop (5’-UACAGGTACCCGAGGC-3’). Individual frames were processed automatically and independently within the EDNA framework82 yielding particularly averaged curves of normalized intensity vs. scattering angle. Frames corresponding to the elution of 0.3 mL min⁻¹ and 3000 frames of 1 were collected using a Pilatus 1M detector (Dectris). The protein of interest were identified in iSPY85, merged and analyzed further using the tools of the ATSAS package84. For A20 C-ter 20 ab initio models were calculated using DAMMIF, averaged, aligned, and compared using DAMAVER. The agreement between scattering curves of E9 and E9exomina/DNA complex and atomic models were calculated using CRYSOL. Curves were plotted with MS Excel.

Data availability. Coordinates and structure factor amplitudes have been deposited in the PDB as entries 5N2E for E9, 5N2G for the Ms2 complex, and 5N2H for the E9exomina mutant. SAXS data and models have been deposited in the SASBDB as entries SASDC5 for the E9exomina/DNA complex, SASDC5N for E9, and SASDCPS for A20 C-ter. Other data supporting the findings of this study are available from the corresponding author on reasonable request.

Received: 13 March 2017 Accepted: 26 September 2017 Published online: 13 November 2017

References
1. Moss, B. in Fields Virology 6th edn, Vol 2 (eds Fields, B. N., Knipe, D. M. & Howley, P. M.). 2129–2159 (Lippincott Williams & Wilkins, Philadelphia, 2013).
2. Challberg, M. D. & Englund, P. T. Purification and properties of the deoxyribonucleic acid polymerase induced by vaccinia virus. J. Biol. Chem. 254, 7812–7819 (1979).
3. Upton, C., Stuart, D. T. & McFadden, G. Identification of a poxvirus gene encoding a uracil DNA glycosylase. Proc. Natl Acad. Sci. USA 90, 4518–4522 (1993).
4. Boyle, K. A., Arps, L. & Traktman, P. Biochemical and genetic analysis of the vaccinia virus uracil DNA glycosylase and A20 proteins as intrinsic components of the DNA polymerase holoenzyme. J. Biol. Chem. 286, 24702–24713 (2011).
5. Czarnecki, M. W. & Traktman, P. Mechanism of antiviral drug resistance of vaccinia virus: identification of residues in the viral DNA polymerase conferring differential resistance to antiviral drugs. J. Virol. 82, 12520–12534 (2008).
6. Durairaj, S. et al. Mutations conferring resistance to viral DNA polymerase inhibitors in camelpox virus give different drug-susceptibility profiles in vaccinia virus. J. Virol. 86, 7310–7325 (2012).
7. Sèle, C. et al. Low-resolution structure of vaccinia virus DNA replication complex. J. Virol. 75, 1087–1098 (1997).
8. Mass, P. I. & Hart, D. J. ESPRIT: a method for defining soluble expression constructs in poorly understood gene sequences. Methods Mol. Biol. Clifton NJ 1586, 45–63 (2017).
9. Fang, W. & Yang, W. Structural Insight into translesion synthesis by DNA Pol II Cell 139, 1279–1289 (2009).
10. Stave, J., Ralke, C., Hennek, G. & Delarue, M. Molecular recognition of canonical and deaminated bases by P. abyssi family B DNA polymerase. J. Mol. Biol. 423, 315–336 (2012).
11. Luo, S. et al. Crystal structure of the herpes simplex virus 1 DNA polymerase. J. Biol. Chem. 281, 18193–18200 (2006).
12. Baranovskiy, A. G. et al. Structural basis for inhibition of DNA replication by aphidicolin. Nucleic Acids Res. 42, 14013–14021 (2014).
13. Wang, J. et al. Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. Cell 89, 1087–1099 (1997).
14. Hogg, M. R. et al. Structural basis for processive DNA synthesis by yeast DNA polymerase ε. Nat. Struct. Mol. Biol. 16, 979–986 (2009).
15. Liu, S. et al. Crystal structure of the herpes simplex virus 1 DNA polymerase. J. Biol. Chem. 281, 18193–18200 (2006).
16. Killelea, T. et al. Probing the interaction of archaeal DNA polymerases with deaminated bases using X-ray crystallography and non-hydrogen bonding isosteric base analogues. Biochemistry 49, 5772–5781 (2010).
17. Kjeldgaard, M., Nissen, P., Thirup, S. & Nyborg, J. The crystal structure of elongation factor EF-Tu from Thermus aquaticus in the GTP-formation complex. Structure 1, 35–50 (1993).
18. Rashid, R. et al. Crystal structure of a Crs5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratoserin congenita. Mol. Cell 21, 249–260 (2006).
19. Shirono, Y. & Steitz, T. A. Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. Cell 99, 155–166 (1999).
20. Zuccola, H. J., Filman, D. J., Coen, D. M. & Hogle, J. M. The crystal structure of an unusual processivity factor, herpes simplex virus UL42, bound to the C terminus of its cognate polymerase. Mol. Cell 5, 267–278 (2000).
21. Archarya, N., Klassen, R., Johnson, R. E., Prakash, L. & Prakash, S. PCNA binding domains in all three subunits of yeast DNA polymerase δ modulate its function in DNA replication. Proc. Natl Acad. Sci. USA 108, 17927–17932 (2011).
22. Jain, R. et al. Structural insights into yeast DNA polymerase delta by small angle X-ray scattering. J. Mol. Biol. 394, 377–382 (2010).
23. DeFilippes, F. M. Site of the base change in the vaccinia virus DNA polymerase gene which confers aphidicolin resistance. J. Virol. 63, 4060–4063 (1989).
45. Magee, W. C., Hostetler, K. Y. & Evans, D. H. Mechanism of inhibition of vaccinia virus DNA polymerase by cidofovir diphosphate. Antimicrob. Agents Chemother. 49, 3153–3162 (2005).

46. Hogg, M., Aller, P., Koningsberg, W., Wallace, S. S. & Doublé, S. Structural and biochemical investigation of the role in proofreading of a beta hairpin loop found in the exonuclease domain of a replicative DNA polymerase of the B family. J. Biol. Chem. 282, 1432–1444 (2007).

47. Barzou, I., et al. Mechanism of concerted RNA-DNA primer synthesis by the human primosome. J. Biol. Chem. 291, 10006–10020 (2016).

48. Doublé, S., Sawaya, M. R. & Ellenberger, T. An open and closed case for all polymerases. Structure 7, R31–R35 (1999).

49. Zahn, K. E., Tchesnokov, E. P., Götte, M. & Doublé, S. Phosphonoformic acid inhibits viral replication by trapping the closed form of the DNA polymerase. J. Biol. Chem. 286, 25246–25255 (2011).

50. Trottitzsch, S., Bienenso, C., Nie, Y., Garzoni, F. & Berger, I. New baculovirus expression tools for recombinant protein complex production. J. Struct. Biol. 172, 45–54 (2010).

51. Kabash, W. XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 (2010).

52. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69, 1204–1214 (2013).

53. Tervilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. 55, 849–861 (1999).

54. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. 62, 1002–1011 (2006).

55. Emseley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).

56. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367 (2011).

57. Ye, Y. & Godzik, A. Flexible structure alignment by chaining aligned fragment pairs allowing twists. Bioinformatics 19 (Suppl 2), ii246–255 (2003).

58. Silvi, A. et al. CATH: a structural and functional annotations for genome sequences. Nucleic Acids Res. 43, D376–D381 (2015).

59. Yumerefendi, H., Tarendeau, F., Mas, P. J. & Hart, D. J. ESPIRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. J. Struct. Biol. 172, 66–74 (2010).

60. van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R. & Groen, G. F. Estimation of protein secondary structure and error analysis from circular dichroism spectra. J. Appl. Crystallogr. 45, 342–350 (2012).

61. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENaCt server. Nucleic Acids Res. 42, W320–W324 (2014).

62. Guermeur, Y., Geourjon, C., Gallinari, P. & Delage, G. Improved performance in protein secondary structure prediction by inhomogeneous score combination. Bioinformatics 15, 413–421 (1999).

63. Kato, S. E. M. et al. Marker rescue mapping of the combined Condit/Dales collection of temperature-sensitive vaccinia virus mutants. Virology 375, 213–222 (2008).

64. Traktman, P., Kelvin, M. & Pacheco, S. Molecular genetic analysis of vaccinia virus DNA polymerase mutants. J. Virol. 63, 841–846 (1989).

Acknowledgements

We acknowledge financing of the project by the French Grants REPLIPOX ANR-13-BSV8-0014 and by research grants from the Service de Santé des Armées and the Délégation Générale pour l’Armement. P.M. acknowledges support from NUPI U (LQ1604). This work used the platforms of the Grenoble Instruct-ERIC Center (ISBG: UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISIBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). Special thanks to F.C.A. Gerard for her help with MALLS-RI and P. Fender for the help with Biacore experiments. We are grateful toward the ESRF for beamtime and thank in particular H. Bethali, A. McCarthy, and M. Brennich for support. We further wish to acknowledge F. Garzoni from the Berger Group at EMBL Grenoble for his kind help with the insect cell culture.

Author contributions

N.T. and F.I. designed research; N.T. and F.I. performed the main experiments with the help of C.D.; P.J.M. and D.J.H. performed ESPRIT on A20. P.M. and E.F. carried out the MS experiments. S.H. performed SAXS measurements and analysis. N.T., C.N.P., W.P.B. and F.I. analyzed data; N.T., W.P.B. and F.I. wrote the paper with input from all the authors.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01542-z.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.