Antiviral Research 87 (2010) 85–94

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Review

The VIZIER project: Overview; expectations; and achievements

Bruno Coutard, Bruno Canard *

Laboratoire Architecture et Fonction des Macromolécules Biologiques, CNRS UMR-6098, Universités Aix-Marseille I et II, ESIL Case 925, 163 Avenue de Luminy, 13288 Marseille, France

ARTICLE INFO

Article history:
Received 12 January 2010
Accepted 14 February 2010

Keywords:
Viral structural genomics
Crystal structure
Emerging virus
RNA virus
Antiviral
Replication

ABSTRACT

VIZIER is an acronym for a research project entitled "Comparative Structural Genomics of Viral Enzymes Involved in Replication" funded by the European Commission between November 1st, 2004 and April 30th, 2009. It involved 25 partners from 12 countries. In this paper, we describe the organization of the project and the culture created by its multidisciplinary essence. We discuss the main thematic sections of the project and the strategy adopted to optimize the integration of various scientific fields into a common objective: to obtain crystal structures of the widest variety of RNA virus replication enzymes documented and validated as potential drug targets. We discuss the thematic sections and their overall organization, their successes and bottlenecks around the protein production pipeline, the "low hanging fruit" strategy, and measures directed to problem solving. We discuss possible future options for such large-scale projects in the area of antiviral drug design. In a series of accompanying papers in *Antiviral Research*, the project and its achievements are presented for each virus family.

Contents

1. Introduction

The VIZIER project was launched October 1, 2004 and ended April 30, 2009. VIZIER is an approximate acronym for 'Viral Enzymes Involved in Replication'. It was the first large-scale, multicountry, co-ordinated research initiative to address the challenging concept of "scientific anticipation" (our own interpretation of these terms will be described later in the text) directed towards getting the necessary scientific knowledge to design drugs against emerging RNA viruses. The funding scheme was a European Commission-funded project referred to as an Integrated Project from the Framework Programme 6 (IP-FP6). It included 25 partners from the EU, listed at the end of this article, with a total EC contribution of 12.9 million Euros (Coutard et al., 2008).

This VIZIER symposium issue presents nine chapters reviewing current knowledge available on most RNA virus order/

* Corresponding author.
E-mail address: bruno.canard@afmb.univ-mrs.fr (B. Canard).

0166-3542/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2010.02.326

families/genus regarding mainly viral genomics, structural biology and their use in drug design, with a special emphasis on what has been achieved within the VIZIER project. Alphaviruses (Gould et al., 2009), coronaviruses (Hilgenfeld et al., this issue), flaviviruses (Bollati et al., 2009), Picornaviruses (Norder et al., this issue), caliciviruses (Rohayem et al., this issue), arenaviruses and bunyaviruses (Charrel et al., this issue), and rhabdoviruses (Assenberg et al., this issue) have their chapters, while ommissions reflect either a deliberate choice at the start of the project (eg., hepatitis C virus, orthomyxoviruses) or at the VIZIER symposium writing stage. For ommitted viruses, although interesting and sometime truly original results have been obtained within VIZIER, the amount of acquired data did not justify a full chapter (eg., reoviruses, astroviruses, paramyxoviruses, ...). In addition, a method chapter is included in the symposium. Indeed, specialized bio-informatics or crystallographic methods have been developed to satisfy VIZIER's special needs. We believe that these methods and tools are difficult to publish in a context other than the VIZIER symposium, but they surely deserve special mention, because they might be useful for the growing structural virology community.

Initially, the core activity of VIZIER was a large-scale viral structural genomics project (SGP) meant to address the problem of the small number of viral protein structures known at the atomic resolution required for antiviral drug design. The SG part *per se* (protein production, crystallization and structure determination) represented ~50% of the overall EC contribution. During the course of the project, the original concept expanded to its present form resting mainly on three scientific fields: (i) a strong "viral genomics" component, (ii) the original "structural genomics" core, and (iii) an integrated drug candidate discovery activity including chemists looping back to basic virology labs for mechanistic and proof-of-concept studies.

Replication enzymes were chosen as primary targets. First, they are essential to the virus life-cycle, and their inhibition leads to effective virus growth suppression. Second, due to their structural and mechanistic conservation across viral families, replication enzymes are less prone to variation than structural proteins. In theory, this should facilitate the subsequent selection of broad-spectrum antiviral molecules, as well as minimise the emergence of drug resistance. Third, viral replication enzymes are often unique or divergent enough from host enzymes to expect decent drug selectivity, minimizing potential toxicity due to cross-inhibition of host enzymes. These enzymes thus constitute so-called "validated targets" in the drug discovery and design field.

The structure-based antiviral drug design era was definitely open with the advent of anti-HIV protease drugs that reached the market in the mid-1990s (for a review, see Mastrolorenzo et al., 2007). On the other hand, there were also an increasing number of enzyme functions inferred from examination of crystal structure, providing new opportunities for rational antiviral drug design. In this context, VIZIER took position in 2004 to increase the connection between the virology/antiviral field and structural genomics.

1.1. The VIZIER concepts

Before the commencement of the VIZIER program, it was already thought that structural genomics could contribute to protein function assignment (Zarembinski et al., 1998). Many large-scale SGPs were launched on targets of full proteomes or subsets of targets from different genomes. The first, which focused on the archaebacteria *Methanobacterium thermoautotrophicum* (Christendat et al., 2000) was followed by other initiatives driven by large consortia in Europe (SPINE, http://www.spineurope.org/), Japan (RIKEN, http://www.rsgi.riken.go.jp/), the USA (PSI, http://kb.psi-structuralgenomics.org/) or transatlantic (SGC, http://www.sgc.utoronto.ca/). The main common objective was to generate structural data related to human health. There were no systematic pan-species approaches, as one genome was most of the time enough to keep a large consortium busy.

In 2004, the VIZIER approach was unique in structural biology. Indeed, the main objective of VIZIER was not only to provide structural data on potential new drug targets, but also, and unlike other consortia, VIZIER was focusing on a very limited range of proteins, those involved in viral replication, that allegedly bore an "antiviral target" potential.

1.2. The pan-viral approach

Existing consortia were relying on the "low hanging fruit" approach. This expression means that the "easy" targets, ie., proteins proceeding without any major difficulty up to structural analysis, were addressed first. In that respect, VIZIER introduced a new variation in the meaning of the expression. For example, only two proteins, the NS3 and NS5 enzymes that are central to replication were considered in the Flavivirus genus. Consequently, a major goal was to obtain at least one crystal structure for any Flavivirus NS3 or NS5, or sub-domains. The term "low hanging fruit" would then refer to the first determined of those, which would in turn help to solve as many other related NS3 or NS5 as possible along the flaviviral phylogenetic tree. The "low hanging" target was therefore considered as a template to accelerate structure determination inside a viral genus or inside a class of enzymes.

This approach initially met with some scepticism amongst the VIZIER crystallographer community despite the frank enthusiasm of the virology community. Indeed, once a first crystal structure is known, structural homologues are usually not as exciting as original unknown proteins bearing the potential to reveal a novel fold. Nevertheless, the pan-viral approach made it possible to hit several birds with one stone! Indeed, viral genomes being small-sized compared to eukaryotic genomes, the pan-viral approach was a "must" to be eligible to a structural genomics denomination (and funding). First, it backed the concept of anticipating emerging viruses (see below). Second, it made it possible to try to produce or crystallize proteins of different isolates, enhancing the chances of success. *A posteriori*, the structural data in complement to the large-scale sequencing (viral genomics) within viral families may help to address the inevitable questions about polymorphisms and drug susceptibility/resistance, as has already occurred with the influenza neuramidase and HIV enzymes.

1.3. The VIZIER "anticipation concept"

The pan-viral approach developed by VIZIER had an immediate consequence. If crystal structures are determined for each virus family, then if a new virus should emerge, its replication enzymes will resemble something that has already been characterized at the structural level, and there might already be an initial drug candidate. In fact, this concept was exemplified by coronavirus research before and during the SARS crisis. In 2002, the Hilgenfeld group (Lübeck University, Germany) and collaborators had determined the TGEV main proteinase structure, the first coronavirus protein crystal structure (Anand et al., 2002). The structural virology field was very young for coronaviruses, in line with most RNA viruses. Interestingly, this structural biology work provided an example of "scientific anticipation" directed towards getting the necessary scientific knowledge to design drugs against emerging RNA viruses. Indeed, TGEV protease inhibitors could be suggested based on the chloromethylketone pharmacophore, due to the active site similarity found with the rhinovirus 3Cpro. When the SARS-CoV emerged, the nucleotide sequence of this new coronavirus was made available on the internet shortly after virus isolation and identification (Marra et al., 2003). It took only days for Dr. Hilgenfeld to figure out

that the SARS protease was closely related to that of TGEV, and to propose both a structural model and that the TGEV inhibitor backbone could serve as a potent starting point for anti-SARS-CoV drug design (Anand et al., 2003). Six months later, the structure of the SARS-CoV main protease structure was solved (Yang et al., 2003).

During the VIZIER project, the chikungunya outbreak in 2005 provided another unexpected opportunity to accelerate the acquisition of knowledge on poorly characterized virus families. The viral *macro*-domains constitute an interesting example. *Macro*-domains are protein domains found in *Coronaviridae*, *Togaviridae*, rubella virus, and hepatitis E virus. The *macro*-domain of coronaviruses has been the subject of several structural studies within and outside VIZIER (Egloff et al., 2006; Saikatendu et al., 2005) and the translation of accumulated knowledge in the *Togaviridae* family undoubtedly accelerated the crystal structure determination of several Alphavirus *macro*-domains, including that of Chikungunya virus for which no crystal structure had been determined before (Malet et al., 2009).

This anticipation concept was communicated repeatedly by the VIZIER consortium (both at scientific meetings and in interactions with the lay press). It has permeated the scientific community, and perhaps to some extent also the general public. It certainly has also stirred debate in the drug design industry, which would not be able to stay clear from the global problem of emerging viruses, and for which anticipation, feasibility, and market evaluation are *sine qua non* conditions for drug development.

1.4. The modular approach

Viral non-structural proteins are usually large polyproteins that comprise several enzyme activities and functions in a single polypeptide chain, which is either later proteolytically processed or not. A possible reason why few of these proteins had been crystallized in 2004 lies in the difficulty of expressing them in a soluble form suitable for crystallization. Drawing on the flaviviral NS5 domain design (Choi et al., 2004; Malet et al., 2007), it was decided to systematize the "modular approach" through the coupling of VIZIER's special bio-informatics and domain prediction, virology, and protein factory. In that sense, VIZIER was at the forefront in the application of high throughput (HTP) techniques to virus research. These techniques were classically present in laboratories performing antiviral screening, but not in academic virology labs.

1.5. The end problem

In the overall process of drug discovery/design, there is a clear tendency to integrate all required scientific fields as much as possible. In VIZIER, defining a starting point was easy, because the need for viral genomics and post-genomics was identified early, as described above. However, one difficulty was to define where VIZIER should stop in the drug discovery/design process and leave further drug development to more specialized developers. One obvious end-line is the unpredictable nature of medicinal chemistry involved in hit-to-lead development. Any discovered hit for which structure-based chemistry has to be implemented would require specialized chemists. It seemed more appropriate to subcontract this secondary chemistry, rather than try to integrate all fields of medicinal chemistry into the VIZIER program. Therefore, VIZIER's aim was to provide structural data for drug design and as an option, hits to begin feeding the drug discovery pipeline.

A pipeline organisation centered on core laboratories from genomics to medicinal chemistry.

In an effort to integrate and connect all of the activities described above, VIZIER relied on its own pipeline from bio-informatics to medicinal chemistry (Coutard et al., 2008). In principle, a pipeline organisation has several advantages. First, documentation of effort is easy and simple, provided that there is an appropriate Laboratory Information Management System (LIMS). Duplication of effort is avoided and the optimal flow and feedback of information are facilitated, as well as sharing protocols and technical improvements through the database server. However, a central pipeline bears the obvious limitation of sequential organisation and dependence. If a single bottleneck appears somewhere in the pipeline, all of the downstream activity groups are at risk of being affected. Ideally, the pipeline has to run at an appropriate steady-state to irrigate various activities. This cannot be pre-existing at time zero of the project, and unpredictable difficulties have expectedly arisen to jeopardize the workflow. These difficulties have actually promoted the organisation of specialized workgroups focusing on targets of interest from a structural and/or antiviral design point of view.

It is now very interesting to evaluate results according to activity fields and virus families (see below), as this will certainly orient future projects. Such achievements in the long term will facilitate the design of molecular structure-based strategies for the inhibition of replication in a broad range of any clinically relevant viruses. In other words, both successes and failures will directly benefit global public health.

2. Results

2.1. Section 1—Bio-informatics and the design of crystallization targets

The bio-informatics section proposes primer sequences, derived from viral sequences, to use in virus-infected cells. Using RT-PCR, these primers generate cDNA target sequences. These 2421 sequences (Fig. 1) are derived either from existing sequences (thus the section work can start at the project start) or from novel sequences derived from the genomics Section 2. Fig. 1 shows that the actual predictions made exceed the objectives of 1500 predictions. Was this objective too low? In fact, most predictors and tools were refined and developed during the project (Gorbalenya et al., this issue). Two tool-bases have been developed or implemented in VIZIER (VaZyMolO and Viralis). They have delivered hundreds of predictions either on a self-serve basis (VaZyMolO) or from request to an expert curator (Viralis). A version of VaZyMolO existed before VIZIER (Ferron et al., 2005), but it was considerably expanded and implemented during the project, and now has free access. The Viralis database was also fully operative after approximately 1 year of the project, allowing a cruise speed of target prediction compatible with VIZIER objectives.

There are a number of specific bio-informatics tools and techniques that have significantly contributed to the crystallization of viral enzymes and domains. Perhaps one of the most important is the identification and elimination of "uncrystallizable" material at an early step. Domain definition had thus the first virtue of avoiding inter- or connecting domains which bring flexibility to a target in solution, impeding crystallization. It was also observed that the expression level of a multi-domain protein is not the algebraic sum of separated domain expression levels. As an example, the N-terminal and C-terminal domains of the flavivirus NS5 express separately to high levels, but the comparative expression of the full-length protein is considerably decreased. The reasons may be multiple (eg., protein degradation and/or cleavage, gene expression Care et al., 2008) and related to flexibility, but the full-length protein, when available, has proven less amenable to crystallization despite intense efforts in several labs.

Among the 43 unique structures, 19 (44%) were obtained using a single predicted construct, whereas for the others, a "multi-

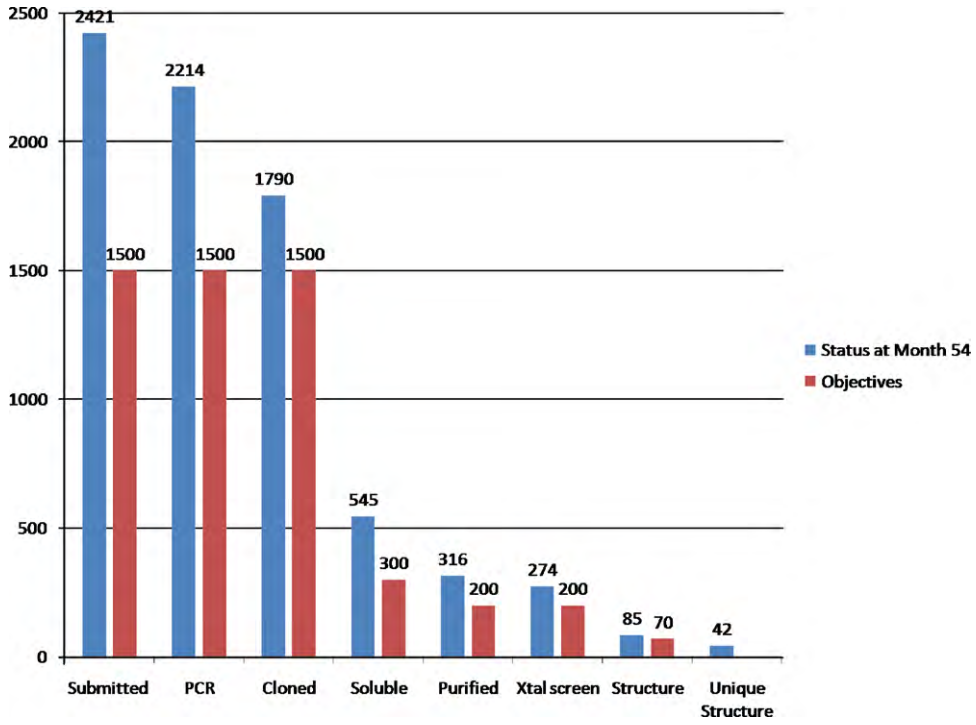

Fig. 1. The status of viral targets through the VIZIER production pipeline. The steps represent the advancement of targets under the form of *in silico* sequences (submitted), cDNA derived from viral RNA (PCR and cloned), recombinant protein expressed under a soluble form (soluble), purified, submitted to crystallogenesis (Xtal screen), and whose crystal structure has been determined (structure). Unique structure refers to a crystal structure of a single protein, either in complex or unliganded. Blue stacks represent the actual achievements compared to the objectives of the project (red stacks) . (For interpretation of the references to color in the citation of this figure, the reader is referred to the web version of the article.)

construct strategy" (see Section 3) was required to improve either solubility or crystallogenesis. Interestingly, amongst the "lowest hanging fruit" targets such as the flavivirus methyltransferases, some homologues crystallized either as a large or as a short form, unpredictably. These results pointed out limitations in the rational sequence-based design. We thus coupled this method to the more empiric multi-boundary construct design, partly explaining the increasing number of domains submitted to the database. Likewise, efforts were made towards a purely random approach to select soluble domains from large target genes refractory to soluble expression after various domain predictions (Cornvik et al., 2005). Although validated by the selection of the soluble fragment of a membrane protein (Manolaridis et al., 2009), this method did not provide new domain boundaries that were missed using bioinformatic predictions.

Among the structures that were determined using a single construct, all were from (+) RNA viruses. Conversely, all soluble and stable domains obtained for (−) RNA viruses were the result of experimental design, using extensive multi-constructs or limited proteolysis on unstable proteins. For (−) ssRNA viral RdRp, bioinformatics were useless to provide boundaries, but could only suggest regions where modules might be found. Three reasons could explain these differences between (+) and (−) ssRNA viruses. First, several structural modules of (+) ssRNA virus were already available in the Protein Data Bank (PDB), rendering domain design easy, whereas no structural data was available on (−) ssRNA viral replicases. Second, most of the (+) ssRNA viral non-structural proteins are processed by a protease suggesting that small modules could behave independently from others. In (−) ssRNA virus non-structural proteins (nsp) no proteolysis occurs and the different domains may need to interact with each other, and may be unstable when produced alone. Nevertheless, successes regarding the influenza RdRp subunits using a random strategy seems to invalidate this hypothesis (Dias et al., 2009; Guilligay et al., 2008). Third, correctly designed modules may fail at the expression step for reasons related to the expression host and not the protein itself.

2.2. Section 2—Viral genomics

The VIZIER consortium gathered a wealth of partial and complete genome sequence data on a very large number of RNA viruses, "the RNA virus genome project". A panel of 252 viruses has been made available within the consortium. In addition, a large number of isolates of clinical interest have been obtained and characterized for "hot" viruses such as dengue virus (>50 isolates), chikungunya virus (>30 isolates), RSV (>50 isolates), human enteroviruses (>500 isolates) or influenza viruses (>50 isolates). Such a reservoir of strains of clinical relevance will be important for future studies. Relatively little and mostly fragmentary information on RNA viral genomes was available at the start of VIZIER in 2004. Complete genome sequence information is essential to obtain a comprehensive picture on genome organisation of RNA viruses, their genetic diversity and variability. However, this ultimate goal was achieved for a few virus families such as:

- Flaviviruses: all tick-borne flaviviruses, all *Aedes*-borne flaviviruses, all *Culex*-borne flaviviruses, all flaviviruses with no known vector (i.e., >40 viruses) were sequenced. In addition, new viruses infecting mosquitoes ("insectiviruses") were identified and 3 new species were completely sequenced (X. De Lamballerie, unpublished). Completion of the genus characterization was reached.
- Arenaviruses: the complete sequencing and analysis of 21 species has been produced. Their phylogeny has recently been reviewed (Charrel, 2008). In addition, virus discovery and genomics yielded characterization of different arenavirus LCMV isolates.

Table 1
Advancement status of 2 cDNA constructs (named "short" and "long") corresponding to the N-terminus of the arenavirus L gene. The cDNAs were amplified from 17 distinct arenaviruses (left column). Each step along the protein production pipeline is indicated, from cloning to crystal structure determination using X-ray crystallography. The status of each target along the pipeline is color coded for each step. Green means that the step was successful, and red that it failed.

Virus	L1 construct	Cloned	Soluble	Purified	Crystal	Structure
Mopeia Virus	Long					
	Short					
LCMV	Long					
	Short					
Guanarito Virus	Long					
	Short					
Pirital Virus	Long					
	Short					
Tamiami Virus	Long					
	Short					
Allpahuayo virus	Long					
	Short					
Bear canyon virus	Long					
	Short					
Amapari Virus	Long					
	Short					
Parana Virus	Long					
	Short					
Machupo Virus	Long					
	Short					
Sabia Virus	Long					
	Short					
Latino virus	Long					
	Short					
Oliveros Virus	Long					
	Short					
Ippy Virus	Long					
	Short					
Lassa virus	Long					
	Short					

- Alphaviruses: 35 complete sequences of chikungunya virus from the Indian Ocean, Gabon, Italy and Madagascar.
- Filoviruses: >20 Ebola virus strains from the CIRMF world's most important collection were be characterized (multigenic sequencing) in collaboration with Dr. Leroy, IRD-Gabon).
- Paramyxoviruses: 35 strains of influenza B virus obtained from 3 different locations (Reims, Marseille, Corsica) have been completely sequenced.
- Bunyaviruses: two new isolates of phlebovirus Toscana virus (full genome sequenced for the newly discovered Massilia virus, the Arbia virus and a strain of Sandfly fever Naples virus).

It is apparent that the viral genomics activity, which was slow at the beginning of the project, later generated more material for the pipeline than it was able to handle!

From a practical point of view, cloning and expressing a gene with many unexpected mutations or polymorphisms reduces the enthusiasm of any motivated scientist who is aiming to publish his results in a scientific paper. This is a notorious limitation for those who work on loosely characterized viral isolates, which have been subjected to many passages in cell culture. The genomics strategy included the need to sequence closely related strains of viruses. This decision was totally vindicated by the observation that closely related strains of the same virus show different characteristics when challenged with potential inhibitors of enzyme activity. These observations also led to the conclusion that future work will require greater emphasis on field isolates of viruses, to avoid the risk of designing antivirals targeting only laboratory strains. Thus, since viruses are amongst the most genetically variable organisms, we now recognise the need to study authentic field isolates whenever possible. This genomics strategy also emphasised the fact that the pre-VIZIER viral world present in the databases as whole genome nucleic acid sequences was wholly inadequate.

Other major beneficiaries of extensive genomics study were the crystallographers who were able to demonstrate a higher than average success rate in solving crystal structures of replicative enzyme domains for a wide range of RNA viruses. Indeed, it might well be that one protein that was resistant to expression or crystallization could be expressed and crystallized from a closely related isolate. In that case, the first crystal structure in an unknown viral family invariably represents a breakthrough. Later on, other related crystal structures are greatly facilitated through molecular replacement techniques making use of the first atomic coordinates. One may think that in terms of visibility and impact, subsequent crystal structures are less interesting for the crystallographer. It was found in many instances, however, that other original structure determination could come from various subsequent targets, expanding interest beyond the first-ever structure determined for a given virus family (Table 1). And from a virologist or epidemiologist point of

Table 2
Comparison of success rates on targets addressed in several structural genomics projects (as reported in the Target DB) *versus* VIZIER. Success rates are indicated for each step along the protein production pipeline up to crystal structure determination.

	Target DB June 10, 2009		VIZIER	
Status	Number of targets	% Success (overall)	Total number of targets	% Success (overall)
Cloned	14,7310	100.0	1790	100.0
Expressed	99,993	67.9	N/A	N/A
Soluble	38,119	25.9	545	30.4
Purified	36,452	24.7	316	17.7
Structures	4599	3.1	85 (43)[a]	4.7 (2.4)[a]

N/A: Only soluble expression of proteins was checked in VIZIER.
[a] Values in parenthesis indicate number of unique structures related to different targets.

view, any structure is potentially interesting, so long as it covers a phylogenetic space or describes an emerging virus.

An unexpected demand from virologists was the wish to be included in the loop when an interesting drug candidate was found at the last stage of the VIZIER pipeline. This logical demand found two main justifications afterward: not only the virology labs were amongst the best qualified partners to assess drug susceptibility spectra on the largest virus collection, but it definitely helped to cement the VIZIER community, by making people from different backgrounds meet and work together.

2.3. Section 3—The protein expression and crystallization factory

Fig. 1 shows that target amplification yielded 2214 cDNAs through RT-PCR from 2421 proposed sequences. In fact, the targets accounting for the difference (2421–2214) were either not addressed or addressed only once. Several expression systems were initially chosen as possible expression systems for viral proteins (Coutard et al., 2008). To that aim, cDNAs were cloned into a set of expression vectors using ligation-independent cloning methods (Walhout et al., 2000; Aslanidis and de Jong, 1990). A library made of 1790 cloned cDNA was built during the project, leading to a large, diverse and well documented resource of viral domains involved in replication. Cloning was a rather efficient step (1790 clones obtained represents a success rate of 80.8%), due to robust ligation-independent procedures that are commercially available. In that respect, large-scale cloning campaigns minimize the handling of freeze–thaw sensitive reagents, perhaps contributing to a rather straightforward and successful step. Noteworthy, the cloning success is evaluated upon sequencing *bona fide* cDNA inserts. Cloning was considered a failure in less than 20% of the cases. Failure was assessed by the absence of any recombinant colonies or the presence of an incorrect cDNA insert (unexpected stop codons, frameshifts, deletions generated by RT-PCR or PCR steps).

During the first 2 years, expression systems were evaluated and *E. coli* was proven to be the most efficient cloning host. It was thus decided to stop exploration of expression methods using non-bacterial expression system, in order to reinforce the production pipeline based on expression in *E. coli* (Berrow et al., 2007; Care et al., 2008). The prokaryotic system became the major protein crystal provider (only one structure determination resulted from a protein produced in insect cells). Two kinds of strategies were successfully used.

2.3.1. The multi-construct strategy

Since most RNA virus enzymes are derived from polyprotein precursors, the VIZIER strategy mainly relied on rational domain (or modules) design (see Section 1 for details). However, in many cases the choice of start and/or stop codons for a module is not obvious, and several gene boundaries must be chosen in a trial-and-error procedure. Combining these multiple starts and stops leads to the cloning of constructs with different lengths for a given domain. As far as crystallization is concerned, this approach, named the "multi-construct strategy", was demonstrated to be about twice as successful as working to fine-tune the soluble expression of a single construct (Graslund et al., 2008). It was used in VIZIER when either a soluble protein was reluctant to crystallization or when poor diffraction data was obtained for a given domain.

One of the first successes using this strategy in VIZIER was the determination of the crystal structure of the West Nile virus RdRp (Malet et al., 2007) using about 30 gene constructs, amongst which only two led to diffracting crystals. Another example provided by Speroni et al. (2009) illustrates the need to screen putative boundaries of the domain, regarding solubility and crystallization. This strategy had also its limitations, in particular for negative-stranded RNA viruses. Indeed, 247 domains were designed based on the L protein of the rhabdoviruses. None of them produced any crystals. Iterative additions or deletions of amino acid sequences may also help protein solubility, as in the case of the echovirus 30 2C domain. The addition of 5 amino acid increments allowed the selection of constructs expressing a soluble protein. Amongst the latter, only one was purified successfully and lead to the determination of a low resolution structure of its hexamer form (Papageorgiou et al., submitted).

2.3.2. Multi-homologue screening

We also took advantage of the biological diversity available in the cDNA library to overcome failures of protein expression, solubility or crystallization. In testing several homologues, we expected that at least one would lead to a soluble protein or crystal. This strategy was initially assessed for flavivirus non-structural proteins. For example, 10 amongst 25 different helicases were produced and purified, most of the methyltransferases were purified, and 10 RdRps were produced for crystallization trials. For more difficult targets (eg., negative-stranded RNA viruses), using a large panel of homologues also helped to provide domains suitable for crystallization. For example, 17 different constructs of the arenavirus L1 domain were cloned and expressed, leading to the purification of two of them (success rate: 18%), highlighting the importance of taking advantage of sequence diversity (see Table 1). It is particularly interesting to note that the Parana virus L1 protein domain led to crystals whose maximum resolution could not be lowered below 8 Å. Coupling both multi-constructs and multi-homologues strategy overcame this difficulty with the LCMV protein which provided crystals diffracting to 2.7 Å.

Fig. 1 shows that the most significant bottleneck remains the expression of soluble protein domains. The success rate of this step is about 30%, whereas expectations were 20% (300 out of 1500), mainly established after examination of the SPINE (European project preceding VIZIER) statistics and data provided by TargetDB (http://targetdb.pdb.org/) (see Table 2). The next step – protein purification – did not use the parallel and semi-automated approach since each protein purification is unique and virtually needs individual care by a dedicated researcher. At this stage, some proteins were purified to homogeneity on the protein platform, and others were transferred to crystallography labs which repeatedly

Table 3
Number of crystal structures determined during the VIZIER project and their distribution within the virus families, orders, and genome types. Only unique structures are reported, ie., a single structure for a single target corresponding to a single virus.

		Unique structures	Sub-total
(+) ss RNA virus	*Flaviridae*	12	35
	Nidovirales	13	
	Togaviridae	3	
	Picornaviridae	4	
	Others	3	
ds RNA virus		3	3
(−) ss RNA virus	Segmented genome	1	5
	Non-segmented genome	4	

prepared protein batches for crystallogenesis experiments. The gap observed here (28% of the proteins remaining unpurified) is again in line with the data provided by the SG consortia (see Table 2).

2.4. Section 4—Crystallization and crystal structure determination of viral enzymes

Only 13% of the purified proteins did not go into crystallization trials. However, Fig. 1 shows that 274 proteins under crystallization trials yielded only 43 unique structures, which were later expanded to 85 structures of protein and protein/ligands, i.e., an average of two crystal structures per unique target (Tables 3 and 4). Although in agreement with the standard results of SG initiatives (see Table 2), this drastic reduction of success (15% of crystallization trials attempted yielded a crystal structure) apparently represents the "Achilles' heel" in structural genomics (Terwilliger et al., 2009). However, there are many observations that may temper and attenuate this apparent failure.

Once the crystal structure of the first protein of its kind or virus family has been determined, interest tends to drop among the crystallographer community, as mentioned above. Therefore, several crystals of the same type of protein (eg., a Picornavirus 3D polymerase, a flavivirus MTase, etc. . .) may well remain unattended for a while, or forever, because their solution is judged not to be worth the challenge! Nevertheless, quality diffraction remains a significant problem. Even though the crystallization drop set-up can be automated, the whole optimization process is labor-intensive and requires full attention from a dedicated investigator, just as classical structural biology projects. Interestingly, some difficult problems were solved by dedicated scientists highly specialized on software for crystal structure determination, as well as crystal structure determination of challenging targets. This was the case for the West Nile RdRp (Malet et al., 2007) and the VPg-3D RdRp Coxsackie virus complex (Gruez et al., 2008), for which a large proportion of amino acids were ill-defined in the structural model, the arenavirus L1 protein domain whose phases could not be easily found, and the coronavirus SUD domain (Tan et al., 2009). The availability of dedicated software whose improved beta versions were made available prior to dissemination in the structural biology community surely helped to solve or improve structural models generated in VIZIER. In addition, a variety of tools were developed to help analysis of the structural models generated in the project (Gorbalenya et al., this issue).

Positive-stranded RNA viruses – and especially flaviviruses and the *Nidovirales* – provided the largest number of structures from the replication complex, as shown in Table 3. For the flaviviruses, the project contributed to the structure determination of all four functional domains of the replication machinery (i.e., the protease, NTPase/helicase, methyltransferase and RNA dependent RNA polymerase), as well as providing additional data regarding the organization of the replication complex. No new fold or original structure was expected in this flavivirus sub-project, but all these targets are essential for viral replication and are thus good candidates for antiviral discovery (further detailed in the following reviews: Sampath and Padmanabhan (2009), Chappell et al. (2008), Dong et al. (2008), Malet et al. (2008)). Each new structure may thus have a direct impact in the antiviral discovery field.

The methyltransferase was the most addressed domain and structural results (22 structures) provided (i) a model for 2′O-methylation (Egloff et al., 2007) as well as N7-guanine methylation (Milani et al., 2009) (ii) an overview of the structural divergence along the flavirirus phylogenetic tree, because at least one structure of one representative of each of the three branches (mosquito-borne, tick-borne, and viruses with no known vector) has been determined. Structure-based drug discovery in complement with experimental validation has already been initiated and has yielded a selection of compounds with IC50 s in the micromolar range (Milani et al., 2009). The structure of the two first flavivirus RdRps (West Nile and dengue virus) have also been determined within the VIZIER project, providing valuable new information for drug discovery in the near future.

In contrast to the flaviviruses, the number and diversity of non-structural proteins and sub-domains is much more important in the family *Nidovirales*. Among the 13 structures determined, all are from Orf1a. No structure was obtained from Orf1b within VIZIER. Nevertheless, the production of soluble proteins of Orf1b enabled functional and biochemical studies (Decroly et al., 2008; Nedialkova et al., 2009). One of the most striking observations that could be made from the structure results concerns the *macro*-domains. At the beginning of the project, the crystal structure of the SARS *macro*-domain was already available (Egloff et al., 2006). Moreover, *macro*-domains were thought to be dispensable for viral replication (Putics et al., 2005) and may be not good candidates for antiviral discovery/design. It could be thus questionable to focus on these domains and their production and structure determination, even though these domains belong to the "low hanging fruit" category. Actually, the structural and functional study of 5 more *macro*-domains showed significant structural differences in substrate affinity and selectivity amongst coronaviruses (Tan et al., 2009; Piotrowski et al., 2009). The invested efforts for coronavirus *macro*-domain design and production were helpful alphavirus *macro*-domain definition and crystallization (Malet et al., 2009). Although many structures are available now but the *macro*-domain role remains elusive, it is certainly much easier to tackle function assessment using reverse genetics in the alphavirus genus than in coronavirus. Indeed, infectious clones or subgenomic replicons are less labor-intensive in the former than in the latter genus. This example illustrates a possible benefit from a pan-virus analysis rather than relying on a unique virus model.

Among the other structures determined for positive-sense, single-stranded RNA viruses, all except one are proteases and RdRp from *Picornaviridae* and *Caliciviridae*. For this exception, the structural work and the associated enzymatic work complemented each

Table 4
Number of crystal structures (unique and unique + liganded = all) determined during the VIZIER project and their distribution within the type of enzyme/protein. RdRp: RNA-dependent RNA polymerase; Methyltransferase: S-Adenosyl-methionine dependent 2′-O or N7-guanine RNA methyltransferase; Helicase: RNA helicase; *macro*-domain: viral protein domain named after the human histone *macro*H2A1.1-domain.

Functional domain	Unique structures	All
RdRp	6	15
Methyltransferase	7	26
Helicase	3	4
Macro-domain	7	15
Protease	7	10
Others	13	15
Total	43	85

other in order to unravel both the first crystal structure for a viral family (*Astroviridae*) and the first demonstration of protease activity (Speroni et al., 2009). In that case, structure and activity of this viral family is clearly lagging behind more studied viruses. As indicated above, one VIZIER concept is to stay aware that any RNA virus may emerge in the near future. In that respect, and although *Astroviridae* work performed in VIZIER could not justify a full chapter in this issue, the human pathogenic astrovirus serine protease crystal structure determined at 2.0 Å resolution may one day speed up drug design. In the case of RdRps, for example, the details of initiation of RNA synthesis are quite diverse in the RNA virus world. Initiation may be primer-independent (eg., flaviviruses), primer-dependent (eg., arenaviruses), protein-primed (eg., picornaviruses), or just unknown (eg., coronaviruses). There are several examples in VIZIER of unexpected results that may have some relevance to drug design. For example, the Coxsackie virus 3D RdRp structure was determined in complex with VPg (Gruez et al., 2008) and found to form a complex different from that of the foot-and-mouth virus 3D RdRp–VPg complex (Ferrer-Orta et al., 2006). Also, the calicivirus RdRp polymerase structure suggested enzyme experiments which showed that caliciviruses can initiate RNA synthesis both in a primer-independent and independent manner (Fullerton et al., 2007).

Three crystal structures have been determined from ds RNA viruses, and among them two are directly involved in virus replication: one RdRp, and one capping enzyme (Sutton et al., 2007). The crystal structure determination of L protein domains involved in the replication of (−) ss RNA viruses was probably the most challenging objective of the project since no structure was available at the start of the project. Several hundreds of L and other protein constructs from a large and diverse virus collection were expressed in *E. coli* and insect cells, contributing to lower the pipeline statistics: only five crystal structures were determined and, amongst them two domains from L proteins. The first one is the N-terminal part of LCMV (referred as (−) ss RNA virus with segmented genome in Table 3). The structural data did not yet relate this domain to any known fold from the PDB and provided only few evidences regarding its function. Nevertheless, reverse genetics is available for arenaviruses (reviewed in de la Torre (2008)), and this could rapidly indicate if this domain carries essential viral functions as well as be a relevant target for structure-based drug design against arenaviruses.

2.5. Section 5—Functional characterization of viral enzymes and their cognate drug candidates

In this section, functions and enzyme activities were tested either on the basis of *in silico* predictions, or more generally when no prediction was available. Once established in one lab, biochemical and enzyme tests were made available to the partnership through the VIZIER website. Indeed, there are a number of generic biochemical tests that can give interesting results aiming at the discovery of novel functions, such as RNA binding, specific nucleotide binding, RNA-cap binding, or co-factor binding.

For example, for arenaviruses and the L1 domains described above, specific GTP binding would point to an RNA-cap snatching enzyme or RNA synthesis priming since an extra G is found at the 5′-terminus of the (−) RNA strand. In that case, selectivity or absence of selectivity towards a methyl group in N7 of the GTP ligand would discriminate between these two functions (Charrel et al., this issue). This information is also sometimes essential to overcome limitations when the protein remains reluctant to crystallization. In that case, the ligand is to be included in crystallization additives. For example, the flavivirus methyltransferases crystallized in the majority of the cases with endogenous methylation co-substrates or co-products, indicating that the presence of the ligand was helpful for stability and crystallization.

When the functional prediction is clear (eg., RdRps), there are a number of assays that can be performed to characterize the enzyme and guide towards future structural studies. The availability of reasonable amounts of pure and soluble protein was also a great incentive to test possible mechanisms for many proteins even if the protein is reluctant to crystallize (eg., the first demonstration of activity for an RdRp in the *Arteriviridae* family Beerens et al. (2007)).

The other and perhaps greater activity of this section is to provide small molecule inhibitors that serve a dual, essential purpose. First, they may become drug leads. The section had an intense activity in mining the literature in search of known inhibitors which could now begin a “second life” because of the newly available crystal structures. Alternatively, the VIZIER crystal structures provided first choice material to understand the mode of action of compounds discovered in partner labs during VIZIER-independent screening campaigns.

Part of the activity of the section was also dedicated to infected cell work, and its connection with structural data. Indeed, selecting and growing mutant viruses resistant to a drug allows mapping with precision a site of action. The availability of structures becomes a significant advantage for drug design. In some cases, data mining identified old drugs that led to renewed interest for a target once precise mapping was achieved (Norder et al., this issue). In the picornavirus case, the discovery of drug candidates preceded high-resolution structural work on the 2C protein (De Palma et al., 2008). The first crystal structure of this essential enzyme will certainly boost the development of drugs targeting this type of enzyme, presumably a SF3 helicase.

3. Conclusion

Over a 4.5-year period, the VIZIER project achieved a close collaboration among scientific fields and cultures that do not usually or naturally meet. It was wonderful to witness how mutual scientific curiosity opened new collaborative actions, while contributing to build a project culture. It is thus not so surprising that the initial structural genomics project shifted towards an antiviral drug development effort, the latter representing an undisputed common ground.

During the project time course, 85 crystal structures were determined (see Fig. 1 and Tables 2–4), exceeding the original objectives for the structural genomics part. Most importantly, the vast majority – if not all – of the crystal structures published have been documented with functional assays. This fact deviates from the idea that such a structural genomics project simply dumps technological data into a common unanalysed resource, as if large-scale nucleotide sequencing would just produce un-annotated sequences. These crystal structures undoubtedly represent valuable information that may or may not be useful for antiviral design in the future, but often represent the first structural scientific knowledge obtained on neglected virus families. The infinite beauty and imagination of nature in protein architecture and enzyme mechanisms will last much longer than a project that has generated enthusiasm and results, and has perhaps contributed to attracting young talents to a multidisciplinary virology science.

Acknowledgements

This work was supported by the VIZIER integrated project (LSHG-CT-2004-511960) of the European Union 6th Framework Programme (FP6). Data were extracted from the VIZIER databases with the great help of Philippe Lieutaud. The authors would like to thank all the VIZIER members for their involvement and enthu-

siasm in this project, and especially: Gorbalenya AE and Snijder EJ (Leiden, the Netherlands); Leontovich AM (Moscow, Russia); Cambillau C, Romette JL, De Lamballerie X and Charrel R (Marseille, France); Gould EA, Owens RJ, Grimes JM and Stuart DI (Oxford, The United Kingdom); Gunther S and Tucker PA (Hamburg Germany); Norder H (Solna, Sweden); Klempa B (Bratislava, Slovakia), Bourhy H (Paris, France); Rohayem J (Dresden, Germany); L'hermite E and Vuillard L (Lausanne, Switzerland); Nordlund P (Stockholm, Sweden); Bolognesi M (Milano, Italy), Mattevi A (Pavia, Italy); Coll M (Barcelona, Spain), Jones TA, Aqvist J and Unge T (Uppsala, Sweden); Hilgenfeld R and Mesters JR (Lübeck, Germany); Bricogne G (Cambridge, The United Kingdom); Neyts J (Leuven, Belgium); La Colla P (Cagliari, Italy); Puerstinger G (Innsbruck, Austria); Gonzalez JP, Leroy E (Nakhonpathom, Thailand and Franceville, Gabon); Langon T, Caprili J and Arnoux S (Lyon, France).

Partner laboratories involved in VIZIER were:

1. Architecture et Fonction des Macromolécules Biologiques, CNRS, and Universités d'Aix-Marseille I et II, UMR 6098, ESIL Case 925, 13288 Marseille cedex 09, France.
2. Structural Genomics Laboratory, IBBMC, CNRS UMR 8619, IFR 115, Université Paris-Sud, Orsay, France.
3. Molecular Virology Laboratory, Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands.
4. Unité des Virus Emergents (EA3292, IFR48, IRD UR034), Faculté de Médecine, 27 boulevard Jean Moulin, 13005 Marseille, France.
5. Center for Ecology and Hydrology-Oxford, Mansfield Road, Oxford OX13SR, UK.
6. Section for hepatitis and enteroviruses, VIV, Swedish Institute for Infectious Disease Control, SE-17182 Solna, Sweden.
7. Institute of Virology, Slovak Academy of Sciences, Dubravska cesta 9, 84505 Bratislava, Slovakia.
8. Laboratoire de la Rage, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France.
9. Bernhard Nocht Institute for Tropical Medicine Bernhard-Nocht-Str. 74, 20359 Hamburg, Germany.
10. The Calicilab, Institut für Virologie, Medizinische Fakultät Carl Gustav Carus, Fiedlerstrasse 42, D-01307 Dresden, Germany.
11. BioXtal, Chemin des Croisettes 22, CH-1066 Epalinges, Switzerland.
12. Oxford Protein Production Facility, The Henry Wellcome Building for Genomic Medicine, Oxford University, Roosevelt Drive, Oxford OX3 7BN, UK.
13. European Molecular Biology Laboratory, Hamburg Outstation, c/o DESY, Notkestrasse 85, 22 603 Hamburg, Germany.
14. Department of Biomolecular Sciences and Biotechnology, CNR-INFM, University of Milano, Via Celoria 26, 20133 Milano, Italy.
15. Department of Genetics and Microbiology, University of Pavia, Via Ferrata 1, 27100 Pavia, Italy.
16. Institut de Biologia Molecular de Barcelona, CSIC, Parc Científic de Barcelona, c/Josep Samitier 1-5, 08028 Barcelona, Spain.
17. Global Phasing, Sheraton House, Castle Park, Cambridge CB3 0AX, United Kingdom.
18. Department of Cell and Molecular Biology, Uppsala University, Biomedical Center, Box 596, SE-751 24 Uppsala, Sweden.
19. Institute of Biochemistry, Center for Structural and Cell Biology In Medicine, University of Lübeck, Ratzeburger Allee 160, Lübeck 23538, Germany.
20. Rega Institute for Medical Research, Minderbroedersstraat 10, B-3000 Leuven, Belgium.
21. Department of Sciences and Biomedical Technologies, University of Cagliari, Cittadella Universitaria SS554, 09142 Monserrato (Cagliari), Italy.
22. IRD-R178 Mahidol University 25/25 Phutthamonthon 4, 73170 Nakhonpathom, Thailand, and IRD-R178 CIRMF, BP 769, Franceville Gabon.
23. Institut für Pharmazie, Abteilung Pharmazeutische Chemie, Universität Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria.
24. Belozersky Institute of Physico-Chemical Biology, Moscow State, University, Moscow 119899, Russia.
25. Department of Biochemistry and Biophysics, Stockholm University, S-106 91 Stockholm, Sweden.

References

Anand, K., Palm, G.J., Mesters, J.R., Siddell, S.G., Ziebuhr, J., Hilgenfeld, R., 2002. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. EMBO J. 21, 3213–3224.

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R., Hilgenfeld, R., 2003. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science 300, 1763–1767.

Aslanidis, C., de Jong, P.J., 1990. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 18, 6069–6074.

Assenberg, R., Delmas, O., Morin, B., Graham, S., De Lamballerie, X., Lauber, C., Coutard, B., Grimes, J. M., Neyts, J., Owens, R. J., Brandt, B., Gorbalenya, A., Tucker, P. A., Stuart, D., Canard, B., Bourhy, H., this issue. Genomics and structure/function studies of Rhabdoviridae proteins involved in replication and transcription. Antiviral Res.

Beerens, N., Selisko, B., Ricagno, S., Imbert, I., van der Zanden, L., Snijder, E.J., Canard, B., 2007. De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. J. Virol. 81, 8384–8395.

Berrow, N.S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N., Stuart, D.I., Owens, R.J., 2007. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. Nucleic Acids Res. 35, e45.

Bollati, M., Alvarez, K., Assenberg, R., Baronti, C., Canard, B., Cook, S., Coutard, B., Decroly, E., de Lamballerie, X., Gould, E.A., Grard, G., Grimes, J.M., Hilgenfeld, R., Jansson, A.M., Malet, H., Mancini, E.J., Mastrangelo, E., Mattevi, A., Milani, M., Moureau, G., Neyts, J., Owens, R.J., Ren, J., Selisko, B., Speroni, S., Steuber, H., Stuart, D.I., Unge, T., Bolognesi, M., 2009. Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. Antiviral Res. 27, Epub ahead of print.

Care, S., Bignon, C., Pelissier, M.C., Blanc, E., Canard, B., Coutard, B., 2008. The translation of recombinant proteins in E. coli can be improved by in silico generating and screening random libraries of a -70/+96 mRNA region with respect to the translation initiation codon. Nucleic Acids Res. 36, e6.

Chappell, K.J., Stoermer, M.J., Fairlie, D.P., Young, P.R., 2008. West Nile Virus NS2B/NS3 protease as an antiviral target. Curr. Med. Chem. 15, 2771–2784.

Charrel, R., Coutard, B., Baronti, C., Canard, B., Nougairede, A., Frangeul, A., Morin, B., Jamal, S., Schmidt, C.L., Hilgenfeld, R., Klempa, B., De Lamballerie, X., this issue. Arenaviruses and Bunyaviruses: from epidemiology and genomics to antivirals. Antiviral Res.

Charrel, R.N., de Lamballerie, X., Emonet, S., 2008. Phylogeny of the genus Arenavirus. Curr. Opin. Microbiol. 11, 362–368.

Choi, K.H., Groarke, J.M., Young, D.C., Rossmann, M.G., Pevear, D.C., Kuhn, R.J., Smith, J.L., 2004. Design, expression, and purification of a Flaviviridae polymerase using a high-throughput approach to facilitate crystal structure determination. Protein Sci. 13, 2685–2692.

Christendat, D., Yee, A., Dharamsi, A., Kluger, Y., Savchenko, A., Cort, J.R., Booth, V., Mackereth, C.D., Saridakis, V., Ekiel, I., Kozlov, G., Maxwell, K.L., Wu, N., McIntosh, L.P., Gehring, K., Kennedy, M.A., Davidson, A.R., Pai, E.F., Gerstein, M., Edwards, A.M., Arrowsmith, C.H., 2000. Structural proteomics of an archaeon. Nat. Struct. Biol. 7, 903–909.

Cornvik, T., Dahlroth, S.L., Magnusdottir, A., Herman, M.D., Knaust, R., Ekberg, M., Nordlund, P., 2005. Colony filtration blot: a new screening method for soluble protein expression in Escherichia coli. Nat. Methods 2, 507–509.

Coutard, B., Gorbalenya, A.E., Snijder, E.J., Leontovich, A.M., Poupon, A., De Lamballerie, X., Charrel, R., Gould, E.A., Gunther, S., Norder, H., Klempa, B., Bourhy, H., Rohayem, J., L'Hermite, E., Nordlund, P., Stuart, D.I., Owens, R.J., Grimes, J.M., Tucker, P.A., Bolognesi, M., Mattevi, A., Coll, M., Jones, T.A., Aqvist, J., Unge, T., Hilgenfeld, R., Bricogne, G., Neyts, J., La Colla, P., Puerstinger, G., Gonzalez, J.P., Leroy, E., Cambillau, C., Romette, J.L., Canard, B., 2008. The VIZIER project: preparedness against pathogenic RNA viruses. Antiviral Res. 78, 37–46.

de la Torre, J.C., 2008. Reverse genetics approaches to combat pathogenic arenaviruses. Antiviral Res. 80, 239–250.

De Palma, A.M., Heggermont, W., Lanke, K., Coutard, B., Bergmann, M., Monforte, A.M., Canard, B., De Clercq, E., Chimirri, A., Purstinger, G., Rohayem, J., van Kuppeveld, F., Neyts, J., 2008. The thiazolobenzimidazole TBZE-029 inhibits enterovirus replication by targeting a short region immediately downstream from motif C in the nonstructural protein 2C. J. Virol. 82, 4720–4730.

Decroly, E., Imbert, I., Coutard, B., Bouvet, M., Selisko, B., Alvarez, K., Gorbalenya, A.E., Snijder, E.J., Canard, B., 2008. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. J. Virol. 82, 8071–8084.

Dias, A., Bouvier, D., Crepin, T., McCarthy, A.A., Hart, D.J., Baudin, F., Cusack, S., Ruigrok, R.W., 2009. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature 458, 914–918.

Dong, H., Zhang, B., Shi, P.Y., 2008. Flavivirus methyltransferase: a novel antiviral target. Antiviral Res. 80, 1–10.

Egloff, M.P., Decroly, E., Malet, H., Selisko, B., Benarroch, D., Ferron, F., Canard, B., 2007. Structural and functional analysis of methylation and 5′-RNA sequence requirements of short capped RNAs by the methyltransferase domain of dengue virus NS5. J. Mol. Biol. 372, 723–736.

Egloff, M.P., Malet, H., Putics, A., Heinonen, M., Dutartre, H., Frangeul, A., Gruez, A., Campanacci, V., Cambillau, C., Ziebuhr, J., Ahola, T., Canard, B., 2006. Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. J. Virol. 80, 8493–8502.

Ferrer-Orta, C., Arias, A., Agudo, R., Perez-Luque, R., Escarmis, C., Domingo, E., Verdaguer, N., 2006. The structure of a protein primer-polymerase complex in the initiation of genome replication. EMBO J. 25, 880–888.

Ferron, F., Rancurel, C., Longhi, S., Cambillau, C., Henrissat, B., Canard, B., 2005. VaZyMolO: a tool to define and classify modularity in viral proteins. J. Gen. Virol. 86, 743–749.

Fullerton, S.W., Blaschke, M., Coutard, B., Gebhardt, J., Gorbalenya, A., Canard, B., Tucker, P.A., Rohayem, J., 2007. Structural and functional characterization of sapovirus RNA-dependent RNA polymerase. J. Virol. 81, 1858–1871.

Gorbalenya, A., Lieutaud, P., Harris, M., Coutard, B., Canard, B., Kleywegt, G.K., Kravchenko, A.A., Samborskiy, D.V., Sidorov, I.A., Leontovich, A.M., Jones, T.A., this issue. Practical application of bioinformatics and informatics by the multidisciplinary VIZIER consortium. Antiviral Res.

Gould, E.A., Coutard, B., Malet, H., Morin, B., Jamal, S., Weaver, S., Gorbalenya, A., Moureau, G., Baronti, C., Delogu, I., Forrester, N., Khasnatinov, M., Gritsun, T., de Lamballerie, X., Canard, B., 2009. Understanding the alphaviruses: Recent research on important emerging pathogens and progress towards their control. Antiviral Res. 16 (July), Epub ahead of print.

Graslund, S., Sagemark, J., Berglund, H., Dahlgren, L.G., Flores, A., Hammarstrom, M., Johansson, I., Kotenyova, T., Nilsson, M., Nordlund, P., Weigelt, J., 2008. The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. Protein Exp. Purif. 58, 210–221.

Gruez, A., Selisko, B., Roberts, M., Bricogne, G., Bussetta, C., Jabafi, I., Coutard, B., De Palma, A.M., Neyts, J., Canard, B., 2008. The crystal structure of coxsackievirus B3 RNA-dependent RNA polymerase in complex with its protein primer VPg confirms the existence of a second VPg binding site on Picornaviridae polymerases. J. Virol. 82, 9577–9590.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R.W., Ortin, J., Hart, D.J., Cusack, S., 2008. The structural basis for cap binding by influenza virus polymerase subunit PB2. Nat. Struct. Mol. Biol. 15, 500–506.

Hilgenfeld, R., Coutard, B., Tucker, P.A., Steuber, H., Canard, B., Snijder, E.J., Gorbalenya, A., this issue. Dissecting the coronavirus proteome in pursuit of strategies to control coronavirus infection. Antiviral Res.

Malet, H., Coutard, B., Jamal, S., Dutartre, H., Papageorgiou, N., Neuvonen, M., Ahola, T., Forrester, N., Gould, E.A., Lafitte, D., Ferron, F., Lescar, J., Gorbalenya, A.E., de Lamballerie, X., Canard, B., 2009. The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. J. Virol. 83, 6534–6545.

Malet, H., Egloff, M.P., Selisko, B., Butcher, R.E., Wright, P.J., Roberts, M., Gruez, A., Sulzenbacher, G., Vonrhein, C., Bricogne, G., Mackenzie, J.M., Khromykh, A.A., Davidson, A.D., Canard, B., 2007. Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. J. Biol. Chem. 282, 10678–10689.

Malet, H., Masse, N., Selisko, B., Romette, J.L., Alvarez, K., Guillemot, J.C., Tolou, H., Yap, T.L., Vasudevan, S., Lescar, J., Canard, B., 2008. The flavivirus polymerase as a target for drug discovery. Antiviral Res. 80, 23–35.

Manolaridis, I., Wojdyla, J.A., Panjikar, S., Snijder, E.J., Gorbalenya, A.E., Berglind, H., Nordlund, P., Coutard, B., Tucker, P.A., 2009. Structure of the C-terminal domain of nsp4 from feline coronavirus. Acta Crystallogr. D Biol. Crystallogr. 65, 839–846.

Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Cloutier, A., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., McDonald, H., Montgomery, S.B., Pandoh, P.K., Petrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G.A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R.C., Krajden, M., Petric, M., Skowronski, D.M., Upton, C., Roper, R.L., 2003. The Genome sequence of the SARS-associated coronavirus. Science 300, 1399–1404.

Mastrolorenzo, A., Rusconi, S., Scozzafava, A., Barbaro, G., Supuran, C.T., 2007. Inhibitors of HIV-1 protease: current state of the art 10 years after their introduction: from antiretroviral drugs to antifungal, antibacterial and antitumor agents based on aspartic protease inhibitors. Curr. Med. Chem. 14, 2734–2748.

Milani, M., Mastrangelo, E., Bollati, M., Selisko, B., Decroly, E., Bouvet, M., Canard, B., Bolognesi, M., 2009. Flaviviral methyltransferase/RNA interaction: structural basis for enzyme inhibition. Antiviral Res. 83, 28–34.

Nedialkova, D.D., Ulferts, R., van den Born, E., Lauber, C., Gorbalenya, A.E., Ziebuhr, J., Snijder, E.J., 2009. Biochemical characterization of arterivirus nonstructural protein 11 reveals the nidovirus-wide conservation of a replicative endoribonuclease. J. Virol. 83, 5671–5682.

Norder, H., De Palma, A.M., Selisko, B., Costenaro, L., Papageorgiou, N., Arnan, C., Coutard, B., Lantez, V., De Lamballerie, X., Baronti, C., Solà, M., Tan, J., Neyts, J., Canard, B., Gorbalenya, A., Coll, M., Hilgenfeld, R., this issue. Picornavirus nonstructural proteins as targets for new anti-virals with broad activity. Antiviral Res.

Piotrowski, Y., Hansen, G., Boomaars-van der Zanden, A.L., Snijder, E.J., Gorbalenya, A.E., Hilgenfeld, R., 2009. Crystal structures of the X-domains of a Group-1 and a Group-3 coronavirus reveal that ADP-ribose-binding may not be a conserved property. Protein Sci. 18, 6–16.

Putics, A., Filipowicz, W., Hall, J., Gorbalenya, A.E., Ziebuhr, J., 2005. ADP-ribose-1″-monophosphatase: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. J. Virol. 79, 12721–12731.

Rohayem, J., Bergmann, M., Gebhardt, J., Gould, E.A., Tucker, P., Mattevi, A., Unge, T., Hilgenfeld, R., Neyts, J., this issue. Antiviral strategies to control calicivirus infections. Antiviral Res.

Saikatendu, K.S., Joseph, J.S., Subramanian, V., Clayton, T., Griffith, M., Moy, K., Velasquez, J., Neuman, B.W., Buchmeier, M.J., Stevens, R.C., Kuhn, P., 2005. Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1″-phosphate dephosphorylation by a conserved domain of nsP3. Structure 13, 1665–1675.

Sampath, A., Padmanabhan, R., 2009. Molecular targets for flavivirus drug discovery. Antiviral Res. 81, 6–15.

Speroni, S., Rohayem, J., Nenci, S., Bonivento, D., Robel, I., Barthel, J., Luzhkov, V.B., Coutard, B., Canard, B., Mattevi, A., 2009. Structural and biochemical analysis of human pathogenic astrovirus serine protease at 2.0 A resolution. J. Mol. Biol. 387, 1137–1152.

Sutton, G., Grimes, J.M., Stuart, D.I., Roy, P., 2007. Bluetongue virus VP4 is an RNA-capping assembly line. Nat. Struct. Mol. Biol. 14, 449–451.

Tan, J., Vonrhein, C., Smart, O.S., Bricogne, G., Bollati, M., Kusov, Y., Hansen, G., Mesters, J.R., Schmidt, C.L., Hilgenfeld, R., 2009. The SARS-unique domain (SUD) of SARS coronavirus contains two macrodomains that bind G-quadruplexes. PLoS Pathog. 5, e1000428.

Terwilliger, T.C., Stuart, D., Yokoyama, S., 2009. Lessons from structural genomics. Annu. Rev. Biophys. 38, 371–383.

Walhout, A.J., Temple, G.F., Brasch, M.A., Hartley, J.L., Lorson, M.A., van den Heuvel, S., Vidal, M., 2000. GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol. 328, 575–592.

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G.F., Anand, K., Bartlam, M., Hilgenfeld, R., Rao, Z., 2003. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc. Natl. Acad. Sci. U.S.A. 100, 13190–13195.

Zarembinski, T.I., Hung, L.W., Mueller-Dieckmann, H.J., Kim, K.K., Yokota, H., Kim, R., Kim, S.H., 1998. Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. Proc. Natl. Acad. Sci. U.S.A. 95, 15189–15193.