



Published in final edited form as:

Nature. 2010 March 11; 464(7286): 167–168. doi:10.1038/464167a.

When four become one

Robert Craigie

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Robert Craigie: bobc@helix.nih.gov

Abstract

Every machine is made of parts. But, as the new structure of the HIV integrase enzyme in complex with viral DNA shows, one could not have predicted from the individual parts just how this machine works.

Retroviruses such as human immunodeficiency virus type 1 (HIV-1) integrate a DNA copy of their genome into their host genome as an obligatory step in their replication cycle. The DNA cutting-and-pasting reactions that lead to integration are mediated by the virus's integrase enzyme. The structures of this enzyme's three domains — the core catalytic domain, the amino-terminal domain and the carboxy-terminal domain — have been solved previously. But the spatial organization of the domains with viral DNA when in the active nucleoprotein complex (known as the intasome), which carries out integration, has remained unknown. In this issue (page 232), Hare *et al.*¹ describe the structure of a retroviral intasome. The structure clarifies many poorly understood aspects of retroviral-DNA integration. What's more, it provides a solid platform for understanding the mechanism by which the inhibitors of this essential viral enzyme act (Fig. 1).

Integration of retroviral DNA has been extensively studied at the biochemical level. Such analyses showed that the DNA cutting-and-pasting steps are mediated by a mechanism similar to that used by the transposase enzymes found in bacteria and more complex organisms. The structures of the catalytic domain of retroviral integrases revealed^{2,3} an active site that is common not only to these transposase enzymes, but also to members of a larger family of polynucleotidyl transferase enzymes⁴. Beyond confirming the mechanism of catalysis, however, these structures were disappointingly uninformative. For instance, they indicated that the catalytic domain is dimeric, with the two active sites on opposite sides of the nearly spherical dimer. This spacing is incompatible with the spacing of the catalysis sites on the two strands of host DNA. Clearly, the functional complex would need at least a pair of such dimers for two active sites to be close to each other.

In addition to the structures of the individual amino- and carboxy-terminal domains of retroviral integrases, the two-domain structures of the catalytic domain with either the amino- or the carboxy-terminal domain have been determined⁵. On the basis of these 'partial' structures, many models have been proposed for the structure of the intasome complex. Hare *et al.*¹ show that all of the earlier models are wrong.

But why has the path to this level of understanding been so tortuous? Retroviral integrases are notoriously difficult proteins to work with: even using mutations that improve their solubility, the integrase and its DNA substrate aggregate under the conditions required for assembly of the active complex. A more fundamental problem is that, on initial binding to viral DNA, the integrase does not show strong sequence specificity, and stable complexes are formed only after the initial binding⁶. Mixing integrase with DNA therefore traps nonspecific aggregates that are unsuitable for structural studies.

Hare *et al.* took advantage of their finding that the integrase enzyme of the prototype foamy virus (PFV) — a close relative of HIV-1 — has much more favourable biophysical properties than its HIV-1 counterpart. Using PFV integrase, they thus successfully crystallized an integrase in complex with viral DNA, a feat that has defied the best efforts of many groups for the past two decades.

As predicted from the biochemical studies, the authors' intasome structure consists of a tetramer of integrase bridging a pair of viral DNA ends. The organization of the integrase tetramer and DNA, however, is unlike any of the postulated models.

Extensive surface contacts between integrase and viral DNA, as well as protein–protein contacts, are the glue holding the structure together. The major contribution of DNA–protein contacts to the stability of the complex is reminiscent of the complex formation between the Tn5 transposase and DNA⁷. The dimer interface of the catalytic domains, which is present in all crystal structures of this domain, is preserved as expected. Only two of the four active sites in the tetramer, one from each catalytic-domain dimer, associate with viral DNA, whereas the amino- and carboxy-terminal domains of these subunits contribute to stabilizing the tetramer. The other catalytic domains in each dimer lie far from the DNA substrate, and their corresponding amino- and carboxy-terminal domains are disordered.

The past few years have seen great progress in the development of therapeutic antiviral drugs that target integrase. These drugs have been developed through large-scale screening efforts. The inhibitor raltegravir — like other promising integrase inhibitors — binds integrase in complex with DNA rather than binding the free enzyme in solution. Knowing little about the integrase structures when in complex with DNA, therefore, it has not been possible to visualize the details of how these drugs interact with the enzymes' active site, much less to seriously consider applying rational design to their development.

Hare *et al.*¹ now change that situation. Their structure succinctly shows how integrase inhibitors displace the 3' end of the viral DNA from the enzyme's active site, thus blocking its integration activity. Furthermore, they find that the active site, which seemed to lie in a featureless landscape in the earlier structures of the catalytic domain alone, is partially buried in a rugged surface — features that may afford potential sites for anchoring inhibitors.

Despite the giant leap forward that this study¹ takes our understanding of retroviral integrases, questions remain. First, four of the domains in the integrase tetramer are disordered. Are they involved in the interaction with the host DNA, as Hare *et al.* suggest? Do they mediate another function, or are they simply redundant? Also, the target host DNA is not present in the structure, although there seems to be only one place for it to go without drastic structural rearrangement.

The stark contrast between reality and the now-redundant previous models cautions against embarking on exuberant model building with few constraints imposed by experimentation. That said, the integrase enzymes of HIV-1 and PFV are sufficiently similar to justify the use of PFV integrase in complex with both viral DNA and antiviral inhibitors for reliably modelling the binding of these inhibitors to the HIV-1 enzyme.

References

1. Hare S, Gupta SS, Volkov E, Engelman A, Cherepanov P. *Nature*. 2010; 464:232–236. [PubMed: 20118915]
2. Dyda F, et al. *Science*. 1994; 266:1981–1986. [PubMed: 7801124]
3. Bujacz G, et al. *J Mol Biol*. 1995; 253:333–346. [PubMed: 7563093]

4. Rice P, Craigie R, Davies DR. *Curr Opin Struct Biol.* 1996; 6:76–83. [PubMed: 8696976]
5. Chiu TK, Davies DR. *Curr Top Med Chem.* 2004; 4:965–977. [PubMed: 15134551]
6. Li M, et al. *EMBO J.* 2006; 25:1295–1304. [PubMed: 16482214]
7. Davies DR, et al. *Science.* 2000; 289:77–85. [PubMed: 10884228]

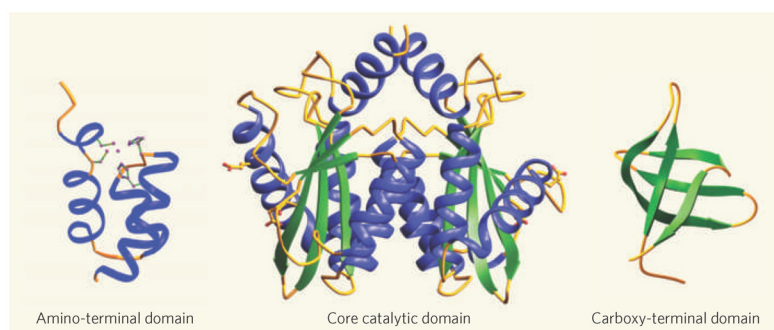


Figure 1. The pieces of the intasome puzzle

An integrase monomer consists of three structural domains: the amino-terminal helical bundle, the core catalytic domain and the carboxy-terminal domain. Viral DNA integration into the host genome is, however, catalysed by a highly stable nucleoprotein complex called the intasome, in which four integrase monomers bridge a pair of viral DNA ends — a structure that Hare *et al.*¹ have now solved, putting the pieces of the puzzle together. Inhibitors of HIV-1 integrase bind to the intasome rather than to free integrase, making the new structure particularly valuable.