Interdomain Communication in Hepatitis C Virus Polymerase Abolished by Small Molecule Inhibitors Bound to a Novel Allosteric Site

The hepatitis C virus (HCV) polymerase is required for replication of the viral genome and is a key target for therapeutic intervention against HCV. We have determined the crystal structures of the HCV polymerase complexed with two indole-based allosteric inhibitors at 2.3- and 2.4-Å resolution. The structures show that these inhibitors bind to a site on the surface of the thumb domain. A carboxylate and phenyl ring substituents, bridged by an indole moiety, fill two closely spaced pockets, whereas a carbonyl substituent forms a salt bridge with an exposed arginine side chain. Interestingly, in the apoenzyme, the inhibitor binding site is occupied by a small a-helix at the tip of the N-terminal loop that connects the fingers and thumb domains. Thus, these molecules inhibit the enzyme by preventing formation of intramolecular contacts between these two domains and consequently precluding their coordinated movements during RNA synthesis. Our structures identify a novel mechanism by which a new class of allosteric inhibitors inhibits the HCV polymerase and open the way to the development of novel antiviral agents against this clinically relevant human pathogen.

The hepatitis C virus (HCV) is a small positive-strand RNA virus responsible for a considerable proportion of acute and chronic hepatitis in humans (1, 2). It is estimated that more than 170 million people worldwide are infected by this virus (3). There is no vaccine available for HCV, and the current therapy, based on interferon and ribavirin, is poorly tolerated and of limited efficacy. Therefore, there are intense research efforts toward the development of new drugs targeting essential HCV enzymes and in particular the polymerase. The HCV nonstructural protein 5B (NS5B) is the RNA-dependent RNA polymerase (RdRp) responsible for replication of the viral genome. NS5B can initiate RNA synthesis by two different mechanisms: primer-independent initiation from the 3′ terminus of the viral genome, also known as de novo initiation (4, 5), and primer-dependent initiation using either DNA or RNA as primers (6). The de novo synthesis is likely used during virus replication in infected cells (7).

The three-dimensional structures of soluble forms of the HCV polymerase genotype 1b, ΔC21 and ΔC55, lacking the last 21 and 55 C-terminal residues, respectively, have been reported (8–10). These structures revealed a classical “right hand” shape formed by the palm, thumb, and fingers domains as initially defined in the Klenow fragment of Escherichia coli DNA polymerase I (11) and showed the presence of an extension in the fingers, the so-called fingertip subdomain, containing two loops, A1 and A2 (secondary structure nomenclature according to Ref. 9), which anchor the fingers to the thumb. As a result, the polymerase has a relatively closed and spherical appearance, and the active site cavity, to which the RNA template and the NTP substrates have access via two positively charged tunnels, is completely encircled. Structural studies have shown that other viral RdRps, such as those of poliovirus (12, 13), human rhinovirus (14, 15), bovine viral diarrhea virus (16), rabbit hemorrhagic disease virus (17), reovirus (18), and bacteriophage φ19 (19), have the same global architecture and a similar closed conformation. Crystal structures of NS5B bound to a short RNA template or to NTPs have also been solved (20, 21). Interestingly, in all these structures, ligand binding occurred without rearrangements in the enzyme domains. Moreover, the structure of the NS5B-GTP complex showed that, in addition to the active site, GTP can also bind to an exposed site on the thumb domain in close proximity to the tip of the fingertip A1 loop (21). This external GTP site, 30 Å away from the polymerase catalytic center, was proposed to exert a regulatory activity by modulating the interactions between the fingers and thumb domains during RNA synthesis. Recently, the structure of the HCV polymerase genotype 2a (this enzyme shares 70% sequence identity with the 1b polymerase) was reported (22). This study revealed two conformations of the enzyme: closed, similar to the apo1b polymerase structure, and a more open structure. Interestingly, in the opened conformation the tip of the fingertip A1 loop, which is a-helical in the closed conformation of the polymerase, moved away from the thumb domain and adopted a β-hairpin-like structure.

The discovery and characterization of a number of structurally diverse non-nucleoside inhibitors (NNIs) of the HCV polymerase have been reported (22–26). These molecules inhibited the polymerase at a stage preceding the elongation step and...
Materials and Methods

Protein Expression and Purification—A variant of HCV NS5B (genotype 1b, strain BK) lacking 55 C-terminal residues (ΔC55) and a P495L point mutant (24) have been expressed and purified as described below. Protein overexpression was performed by induction of mid-log phase BL21 DE3 E. coli cells with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside for 22 h at 18 °C in LB medium. Harvested bacteria were resuspended in pre-cold lysis buffer (20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 20% glycerol, 0.2% β-mercaptoethanol, 1 mM EDTA, 10 mM dithiothreitol (DTT), and Complete EDTA-free protease inhibitor mixture tablets) and lysed with a cell disruptor (Costant System). Lysate was incubated with DNAse I (5 units/ml) in the presence of 10 mM MgCl2 and clarified by centrifugation at 35,000 × g for 1 h. The supernatant was loaded onto a heparin column, pre-equilibrated with A-HeP (20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 20% glycerol, 0.2% β-mercaptoethanol, 1 mM EDTA, 10 mM DTT) at 4 °C, and the protein was eluted with a salt gradient in 700 mM NaCl. Protein fractions were pooled and diluted with the A-HeP buffer without NaCl to reach 150 mM NaCl and loaded by centrifugation at 35,000 g. Crystals were then transferred for 1 min to a cryoprotectant solution containing 100 mM MES, pH 6.0, 14% polyethylene glycol 8,000, 14% 2-propanol, 10 mM DTT, 10 mM MnCl2, and 1.25 mM compound 1 or 2. Diffraction data were collected at 100 K from cryo-cooled crystals at the ID14-H1 beam-line at the European Synchrotron Radiation Facility (Grenoble, France). Data were processed with MOSFLM and scaled with SCALa (31). Data collection statistics are summarized in supplemental Table II. Further data processing and analysis were performed with the CCP4 suite of programs (31).

Structure Determination and Refinement—Crystals of the apoprotein grew in space group P212121, with two molecules in the asymmetric unit (9), but after soaking, the space group shifted to P212121 with one molecule in the asymmetric unit (supplemental Table II), similar to what was described for the structure of the polymerase-GTP complex (21). The two structures of the ΔC55 polymerase-inhibitor complexes were solved by the molecular replacement method with the program AMoRe from CCP4 (31) using the structure of the apoΔC55 as a search model (9). Model building with REFMAC from CCP4 (31), using a maximum likelihood target, resulted in the final models (supplemental Table II). No electron density was present for residues 22–35, 148–152, and 532–536, which were therefore excluded from the refinement. Both structures satisfy all the criteria set by the program PROCHECK (32), and all amino acids lie in allowed regions of the Ramachandran plot.

Results and Discussion

Inhibition of HCV Polymerase—The key pharmacophore elements of the allosteric inhibitors used in this study include a central scaffold of a 6,5-fused ring system, which proved to be optimal (27), an acid, a cyclohexyl and an aryl substituent (Table I). The only difference between the two inhibitors resides in the indole N-aryl substituent. Both compounds inhibited the purified enzyme with IC50 values in the low nM range and subgenomic HCV RNA replication in the sub-μM range (Table I). Inhibition was independent of the form of recombinant polymerase used as both full-length and C-terminally truncated ΔC21 and ΔC55 were inhibited with similar IC50 values (for compound 2 the IC50 values for the full-length and ΔC21 enzymes were 22 and 100 nM, respectively).

Inhibitor Binding and Polymerase Conformational Change—The crystal structures of the ΔC55 polymerase in complex with compound 1 (2.3-Å resolution) and compound 2 (2.4-Å resolution) were obtained by soaking procedures. In both cases, anal-
analysis of the initial electron density map showed clearly defined electron density for the inhibitor (supplemental Fig. 6). However, different from what was predicted (24), binding did not occur at the exposed GTP binding site but instead 8 Å from it, in a site occupied in the apoenzyme by the small α-helix A at the tip of the fingertip A1 loop (Fig. 1). This loop anchors the fingers to the thumb in the apoenzyme and opens in the direction of the solvent, becoming partially disordered in the inhibitor-bound structures with no electron density present for residues 22–35 (Fig. 1). Electron density was also missing for residues 148–152, part of the fingertip A2 loop that is adjacent to the A1 loop and also connects the fingers and thumb (Fig. 1). This region is already flexible in the ΔC55 apoprotein (9) but is completely disordered in the inhibited structures, presumably because of the lack of stabilization provided in the apoenzyme by the interaction between A2 Val-147 and A1 Met-36. Another consequence of the displacement of α-helix A and the resulting weaker interaction between the thumb and fingers is a slight opening of the polymerase. Indeed, when the palm and thumb domains of the inhibited structure and the apoprotein are superimposed, we observed a rigid body rotation of the fingers of 2–4° in a clockwise direction with respect to the palm. Furthermore, the differences in Cu positions in the fingers domain varied between 0.5 and 2.1 Å (Fig. 1B). It cannot be excluded that the rearrangement of the polymerase domains may have been limited in the crystals by packing interactions and thus could be even larger in solution. Indeed, crystals could be soaked in the presence of the inhibitor only for a relatively short amount of time. An opening of the polymerase, following inhibitor binding, might prevent either the correct positioning of RNA and NTPs required for catalysis or the translocation of the template and product.

Inhibitor-Enzyme Interactions—Despite having a different indole N-acetyl substituent (Table I) the two inhibitors bind almost identically to the polymerase (Fig. 2A). In fact only two carbon atoms of the N-acetyl substituent are in van der Waals contact distance with the backbone carbonyl oxygen atoms of Leu-492 and Gly-493 (Fig. 2B), explaining the similar binding modes and IC₅₀ values for the two molecules (Table I).

The inhibitor-polymerase interaction is mainly hydrophobic. The only exception is a salt bridge formed by the carboxylate group of the inhibitor and the guanidinium group of Arg-503. Plausibly, Arg-503 and the proximal His-428 side chains create a positive electrostatic potential favorable for the binding of acidic ligands (Fig. 2B). The cyclohexyl and the phenyl ring substituents likely contribute to most of the binding energy by filling two closely spaced pockets of the binding site. The side chains of Leu-392, Ala-395, Thr-399, Ile-424, Leu-425, His-428, and Phe-429 form a deep pocket, filled almost completely by the cyclohexyl group of the inhibitor, whereas the Val-37, Leu-392, Ala-393, Ala-396, Leu-492, and Val-494 side chains form a narrower pocket occupied by the inhibitor phenyl ring (Fig. 2B). Interestingly, none of these residues changed conformation compared with the apoprotein. The indole core bridges the two groups and also contributes both in filling the pockets and in the interaction with the protein. One side of the indole moiety points toward the pocket interior and interacts with the Val-494 and Trp-500 side chains, whereas the other side makes van der Waals contacts with the exposed Pro-495 pyrrol ring on the external lower side of the pocket (Fig. 2B). The latter interaction likely explains the lack of inhibition by this class of compounds of the NS5B P495L point mutant (Table I) (24).

The structures suggest several ways to improve the potency of this class of inhibitors (Fig. 2B). In the direction perpendicular to the plane of the indole moiety the occupancy of the pocket is not complete, and slightly bulkier moieties may increase the contact area by approaching Val-494. Substituents inserted at the para-position of the phenyl ring could extend into a protein hydrophobic cleft (designated as “a” in Fig. 2B), significantly increasing the surface contact area. The N-acetamide substituent could be modified to form hydrogen bonds with the backbone oxygen of Leu-492 and Gly-493. Finally, a larger group could substitute for one of the carbohydrate oxygens, so as to extend the inhibitor into a second hydrophobic region (designated as “b” in Fig. 2B), while keeping with the other oxygen the polar interaction with the side chain of Arg-503 (Fig. 2, A and B). In agreement with this hypothesis, a structurally related indole-scaffolded inhibitor, containing the same cyclohexyl and phenyl substituents and an additional elongated amide extension in this position, inhibits the NS5B polymerase with a subnanomolar IC₅₀.

Indole-based NNIs and Other Small Molecule Binding Sites—Comparison of the NS5B-GTP complex structure (21) with our structures shows that the binding sites for GTP and for the indole-based inhibitors are close in space but clearly distinct. The GTP molecule lies on a shallow surface in proximity to the fingertip α-helix A, whereas the inhibitor fills the pocket occupied in the apoenzyme by the same α-helix (Fig. 2, A and C). Moreover, GTP binding is not associated with any

---

2 S. Harper, unpublished data.
protein conformational change (21), whereas inhibitor binding requires a displacement of the α-helix A and part of the A1 loop (Fig. 2, C and D). However, GTP and inhibitors make some common interactions with polymerase atoms (Fig. 2, B and C). Both molecules are at van der Waals contact distances with Pro-495, although they lie on opposite sides of the pyrrol ring. Further, both molecules interact with the Arg-503 side chain; the GTP molecule establishes one water-mediated hydrogen bond, whereas the inhibitor forms a salt bridge via its carboxylate group (Fig. 2, B and C). Interestingly, in the NSSB-GTP complex the ribose moiety of GTP hydrogen bonds the guanidinium group of Arg-32 (Fig. 2C). This interaction may contribute to tightening the association of the fingertip A1 loop with the thumb and could disfavor the opening of the loop and consequently the access of the inhibitor to its binding site. Together, the interaction with common residues and between Arg-32 and GTP provides an explanation for the antagonistic effect that high concentrations of GTP have on polymerase inhibition by this class of compounds (24).

The binding site for the indole-based inhibitors is also distinct from the binding site for another group of NNIs based on thiophene, phenylalanine, or dihydropyranone scaffolds. These compounds bind to a common site on an elongated hydrophobic cleft in the thumb domain near the C terminus of the polymerase (Fig. 3A) (21–23, 26). Although the binding sites for this group of compounds and for the indole-based inhibitors are 14 Å apart, they reside on opposite sides of the thumb domain helix Q, which in turn contacts helices S and T (Fig. 3B). Thus, it is conceivable that these two distinct inhibitor binding sites are functionally linked through these secondary structural elements. This hypothesis is supported by recent structures of the 2a polymerase alone and bound to two thiophene-based NNIs (22). In this study the tip of the 2a polymerase A1 loop was shown to adopt two different conformations, helical and β-hairpin-like. The α-helical conformation was associated with a closed polymerase structure similar to the 1b apostucture, whereas in the β-hairpin-like conformation the polymerase adopted a more opened structure (22). Interestingly, the inhibitors could bind only to the closed form of the enzyme and induced a transition to the opened conformation. Thus, although this study did not explain how inhibitor binding would trigger these conformational changes, the results suggest that these inhibitors may act by causing a perturbation of the interaction between the fingertip A1 loop and the thumb domain and simultaneously locking the enzyme in an opened conformation.
“inactive” conformation (22). Therefore, even if these two classes of NNIs bind to sites located in different regions of the thumb, they may ultimately inhibit the polymerase by a common mechanism.

The Role of the Fingertip α-Helix A—The structures of the inhibited polymerase complexes reported here imply that in the apoenzyme the A1 loop is conformationally flexible allowing access of the indole-based inhibitors to the otherwise buried binding site. In the apostructure, the Leu-30 and Leu-31 side chains from α-helix A only partially fill the pockets occupied by the inhibitor cyclohexyl and phenyl substituents, respectively (Fig. 2D). In fact, not only are the two side chains smaller than the two inhibitor cyclic substituents, but also, because of constraints imposed by the α-helical structure in this region of the A1 loop, they protrude less deeply into the corresponding pockets. It is attractive to speculate that a suboptimal anchoring of fingers to the thumb established by α-helix A might provide the polymerase with the conformational flexibility required during the different steps of RNA synthesis. The flexibility of the A1 loop and its role in holding the thumb and fingers together are further supported by the recently solved structures of the polymerase genotype 2a (22) (see above).

The high sequence conservation among HCV genotypes of the region displaced by the inhibitor, as well as of residues forming the inhibitor binding pockets (Fig. 4), supports an essential function of α-helix A-thumb domain interaction for NS5B activity. Consistently, mutation of Leu-30 to Ser or Arg results in an inactive enzyme (33). Helical regions at the tip of fingertip loops connecting the fingers and thumb are present in other viral RdRps (12–17, 19). The similarities in the structure and enzymatic mechanism (34) between NS5B and these polymerases suggest that flexibility of the fingertip loops could also be a feature of some of these RdRps, and inhibitor binding sites similar to the one described here could also be available for them.

Conclusion—Previous biochemical studies have suggested that the HCV polymerase catalyzes RNA synthesis through an ordered stepwise mechanism (24). The polymerase-RNA template complex is first formed rapidly. Second, a productive enzyme elongation complex is formed more slowly (24). The
indole-based inhibitors, similar to what has been observed for the structurally related benzimidazole-based inhibitors (24), bind and inhibit NS5B only prior to formation of the latter complex.3

Several pieces of evidence support our structure as a genuine model for the high affinity indole-based inhibitor polymerase interaction. First, the competition between these inhibitors and GTP and the lack of inhibitor binding of the point mutant P495L are in agreement with the structures here reported. Second, this class of inhibitors bind to the apoenzyme and to the initial template-enzyme complexes with similar affinities (24) suggesting that they bind similarly to both forms of the polymerase. Third, the almost perfect fit of the cyclohexyl and phenyl rings into the α-helix A-derived binding site pockets is suggestive of a high affinity interaction. Fourth, although our inhibited structure revealed a rigid body rotation of the fingers relative to the palm and thumb domains, we did not observe any changes in the structure of the thumb domain itself. This suggests that the details of the interaction between thumb and inhibitor would be preserved even if in solution the polymerase adopted a more opened conformation when bound to the inhibitor.

When the published biochemical data are combined with the structures reported here, a model for the NS5B inhibition by the indole-based compounds can be proposed (Fig. 5). Before formation of a productive initiation complex, the fingertip α-helix A is flexible and opens toward the solvent allowing binding of the inhibitor. After binding, the connection between the fingers and thumb domains is, at least partially, abrogated, and the Λ1 loop can no longer mediate the concerted rearrangement of the polymerase domains needed for formation of a productive RNA-enzyme complex. Interestingly, when added after formation of the productive complex, the inhibitor did not bind or inhibit (24, 25) suggesting that during elongation either the binding site is not available because of a tighter interaction between the Λ1 fingertip residues and thumb or because the polymerase has assumed a somewhat different shape.

In conclusion the structures reported here have identified a novel allosteric inhibitor binding site on the HCV NS5B polymerase that could not be predicted based on previous work and revealed the mechanism of inhibition of a new class of HCV NNIs. Moreover, the structures have shown the details of the interactions established by the inhibitor suggesting ways to improve the potency for this class of molecules. Finally, our study further supports a model whereby the polymerase Λ1 loop regulates the coordinated movements of the fingers and thumb during the polymerase reaction cycle.

Acknowledgments—We thank the staff at the European Synchrotron Radiation Facility for assistance in the data collection, David Olsen, Alessandro Vannini, and Edwin Rydberg for useful discussions, and Emanuela Emili for help with the artwork.

REFERENCES


3 L. Tomei, unpublished results.