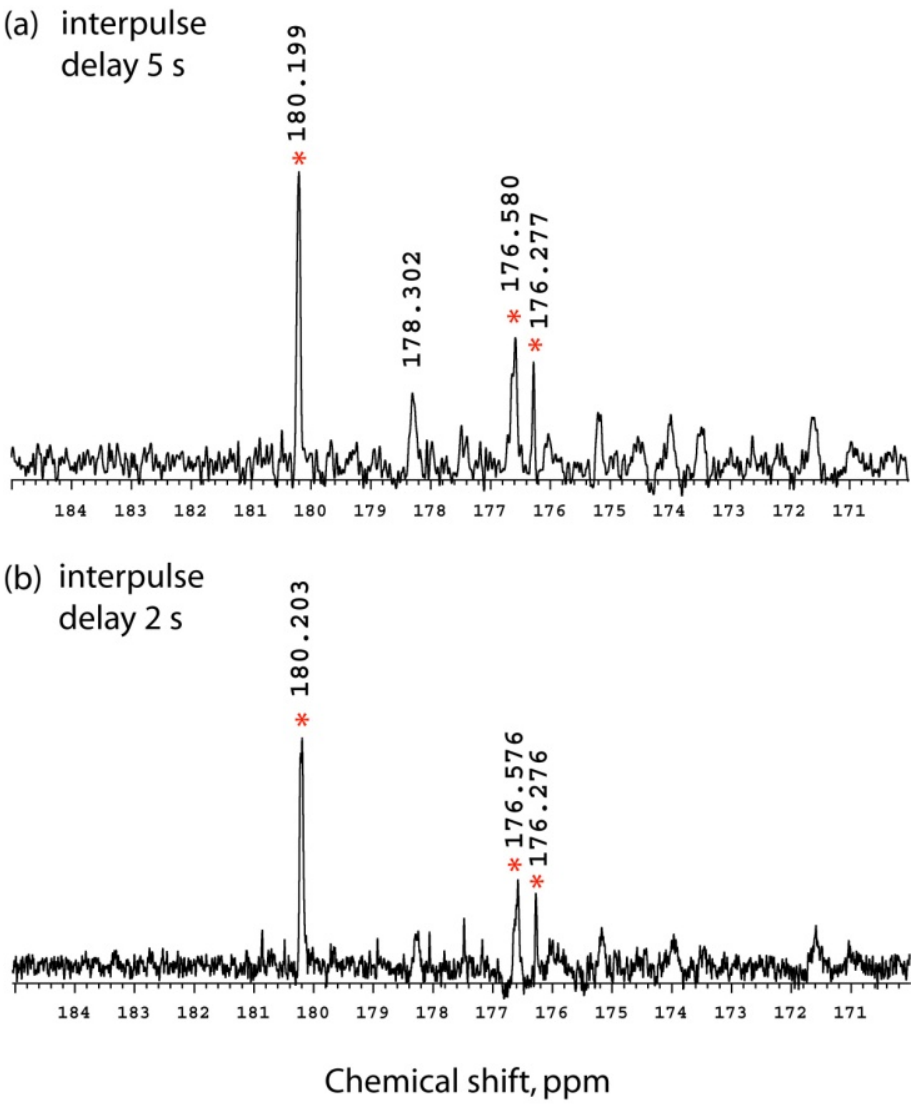
Supplementary Information

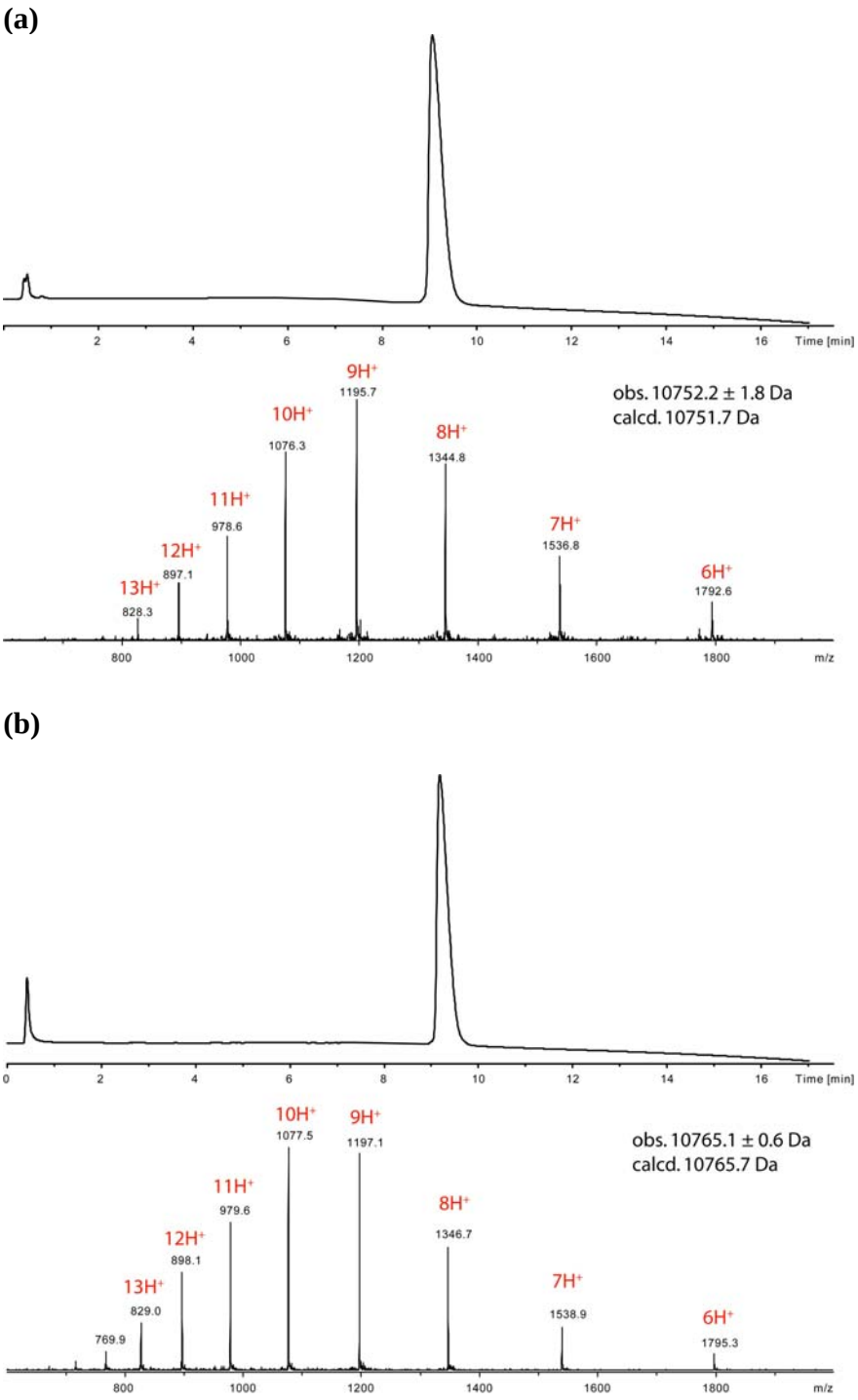
Ionization state of the catalytic dyad Asp25/25' in the HIV-1 protease: NMR studies of site-specifically ^{13}C labeled HIV-1 protease prepared by total chemical synthesis.

*Vladimir Yu. Torbeev, Stephen B. H. Kent**

^{1}H-decoupled ^{13}C-NMR spectra were acquired on Varian Unity Inova 900 spectrometer operating at 226.3 MHz for ^{13}C-nucleus. For unliganded samples, typically ~5000 transients were collected with acquisition time 0.5 s and inter-pulse delay 1 s; for samples with inhibitors, 5000 – 10000 transients were acquired with acquisition time 0.5 s and inter-pulse delay varied from 2 s to 5 s (see legends of Fig. 2 and Supplementary Fig. 1 in the article). All samples were prepared in 18.9 mM Na.phosphate buffer (pH 5.7), containing 5.4 % (v/v) D_2O and 100 μM DSS-d_6. Inhibitors (MVT101 and KVS-1) were added in 4-fold molar excess to the solution of [*L*-Ala51/51']HIV-1 protease after dialysis. Concentrations of protein were determined by integration of total LC-peak at 280 nm and were 0.34 mM for unliganded [*L*-Ala51/51']HIV-1 protease, 0.41 mM for [*D*-Ala51/51']HIV-1 protease, 0.29 mM for [Aib51/51']HIV-1 protease, 0.4 mM for the complex of [*L*-Ala51/51']HIV-1 protease with MVT-101 inhibitor, and 0.22 mM for the complex of [*L*-Ala51/51']HIV-1 protease with KVS-1 inhibitor. ^{13}C chemical shifts were referenced indirectly to DSS-d_6 using γ_C/γ_H ratio.[1] For all experiments the temperature was set to 3.3 ºC to slow down autoproteolysis. (1,4)-^{13}C-Aspartic acid was purchased from Cambridge Isotopes, and was then side-chain-protected with allyl group[3] and Boc-protected at the alpha amino group, and incorporated at Asp25 into the appropriate synthetic peptide segment as described.[2] Site-specific ^{13}C-labelling was performed by total chemical protein synthesis, as previously described.[2] See Supplementary Figure 2 and Supplementary Figure 3 for LC-MS analytical characterization of the chemically synthesized full-length [1-99] HIV-1 protease wild-type and analogue polypeptide chains.

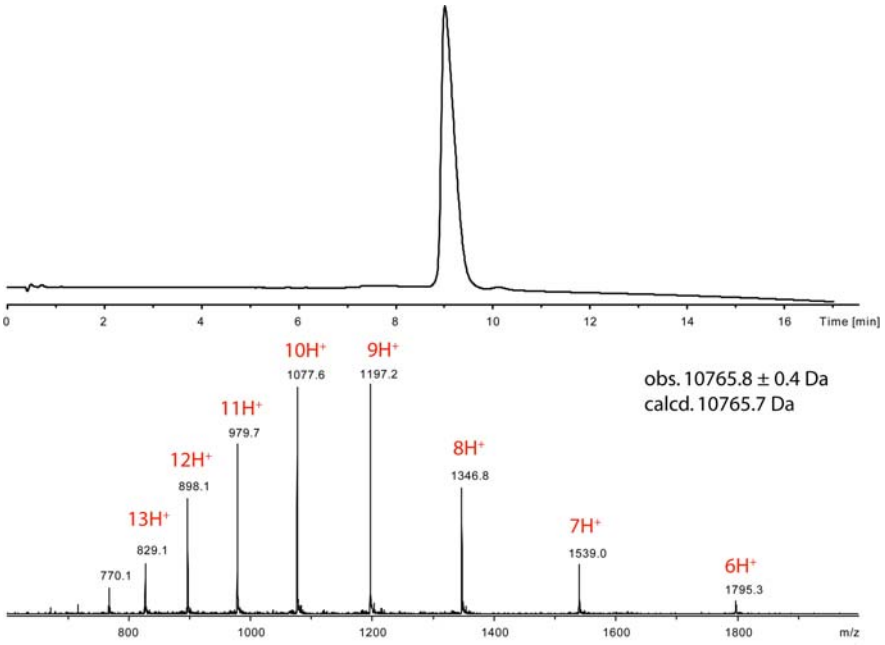


Supplementary Figure 1. 226 MHz ^{13}C-{^{1}H} NMR spectra for complex of [*L*-Ala51/51']HIV-1 protease with the reduced isostere inhibitor MVT-101, acquired with inter-pulse delay of 5 s (in a) and 2 s (in b). Red asterisk indicates peaks originating from unfolded peptidic autoproteolysis products. In (b) spectrum is dominated by autoproteolytic product and signals originating from protein are relaxation-filtered by application of the shorter inter-pulse delay.

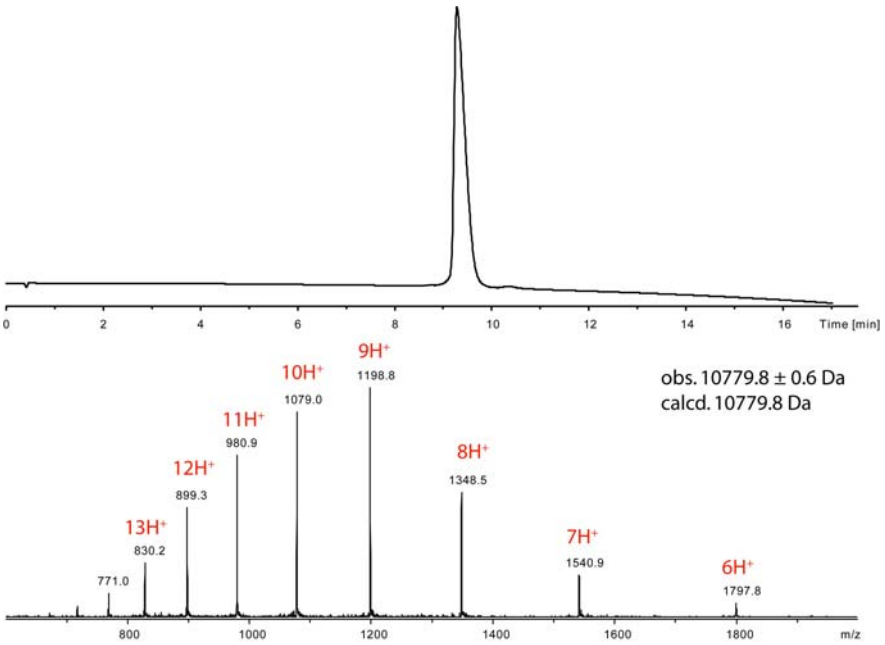


Supplementary Figure S2. Analytical HPLC (λ = 214 nm) and ESI mass-spectra of ^{13}C-labeled (a) wild-type [1-99]HIV-1 protease, and (b) *L*-Ala51 [1-99]HIV-1 protease.

(a)



(b)



Supplementary Figure 3. Analytical HPLC (λ = 214 nm) and ESI mass-spectra of ^{13}C-labeled (a) *D*-Ala51 [1-99]HIV-1 protease, and (b) Aib51 [1-99]HIV-1 protease.

References:

1. Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995***, 6*, 135 – 140.

2. Torbeev, V. Y.; Mandal, K.; Terechko, V. A.; Kent, S. B. H. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4554-4557.

3. Belshaw, P. J.; Mzengeza, S.; Lajoie, G. A. *Synth. Commun.* **1990**, *20*, 3157 – 3160.