Published in final edited form as:

Chembiochem. 2020 November 02; 21(21): 3051–3055. doi:10.1002/cbic.202000263.

Probing the interaction between HIV-1 protease and the homodimeric p66/p66' reverse transcriptase precursor by double electron-electron resonance EPR spectroscopy

Thomas Schmidt, John M. Louis, G. Marius Clore*

Laboratory of Chemical Physics, National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520 (USA)

Abstract

Following excision from the Gag-Pol polyprotein, HIV-1 reverse transcriptase is released as an asymmetric homodimer comprising two structurally dissimilar but amino acid sequence identical p66 subunits. Subsequent cleavage of the RNase H domain from only one of the subunits, denoted as p66', results in the formation of the mature p66/p51 enzyme in which catalytic activity resides in the p66 subunit and the p51 subunit (derived from p66') provides a supporting structural scaffold. Here we probe the interaction of the p66/p66' asymmetric reverse transcriptase precursor with HIV-1 protease by pulsed Q-band double electron-electron resonance EPR spectroscopy to measure distances between nitroxide labels introduced at surface engineered cysteine residues. These data suggest that the flexible, exposed linker between the RNaseH and connection domains in the open state of the p66' subunit binds to the active site of protease in a configuration that is similar to that of extended peptide substrates.

Graphical Abstract

Q-band pulsed double electron-electron resonance EPR spectroscopy shows that binding of HIV-1 protease to the asymmetric p66/p66' homodimer precursor of HIV-1 reverse transcriptase favors an open conformation of the p66' subunit allowing the linker connecting the RNaseH and connection domains to bind to the active site of protease in an extended conformation.

Keywords

HIV-1 protease; HIV-1 reverse transcriptase precursor; deuteration; EPR; DEER

[*] mariusc@mail.nih.gov.

Supporting information for this article is given via a link at the end of the document.

HIV-1 reverse transcriptase (RT) is responsible for catalysing the conversion of single stranded viral RNA into proviral DNA for integration into the host genome, and consequently is the target for several drugs that play a major role in the treatment of AIDS.^[1] The p66 subunit of RT is initially expressed within the Gag/Pol polyprotein and released, following proteolytic cleavage by HIV-1 protease (PR), either as a p66 monomer that subsequently dimerizes or a pre-formed p66 homodimer.^[2–3] Further proteolytic attack by PR cleaves only a single RNase H domain from the p66 homodimer precursor to yield the mature p66/p51 heterodimer where the two subunits, p66 and p51, differ by the presence or absence, respectively, of a RNase H domain (Fig. 1).^[2–5] Although the remaining four domains (finger, palm, thumb and connection) of RT are shared between p66 and p51, their relative orientations in the two subunits are entirely different.^[6] Further, catalytic activity and the binding site for RT inhibitors reside exclusively within the p66 subunit, while p51 functions as a structural scaffold to position the domains of p66.^[7] We recently studied the p66 homodimer precursor by double electron-electron resonance (DEER) EPR spectroscopy,^[8] a technique for measuring distances between pairs of nitroxide spin labels ranging from 15 to ~60 Å for proteins at natural isotopic abundance and up to ~160 Å for fully perdeuterated proteins.^[9–11] We showed that the p66 homodimer (hereafter referred to as p66/p66') displays exactly the same asymmetric arrangement of subunits as the mature p66/p51 RT and that only the RNase H domain of the p66' subunit is accessible to cleavage by PR (Fig. 1).^[8] The RNaseH domain of p66' is connected to the core of p66' (made up of the thumb, finger, palm and connection domains) by a flexible linker and samples two major conformers: a compact state in which the RNaseH domain of p66' folds back onto the p66 subunit; and an open state in which both the linker and the RNaseH domain are fully solvent accessible and sample a broad ensemble of conformers.^[8] Here we use DEER to directly probe the interaction of PR with the p66/p66' RT precursor.

Both p66 and PR (bearing an active site D25N mutation to abolish catalytic activity) were fully deuterated and nitroxide (R1) spin-labeled with S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate at engineered surface exposed cysteines. DEER experiments were conducted at Q-band using a conventional 4-pulse DEER sequence.^[12] $P(r)$ distributions were derived from the DEER echo curves using three different methods: the program DD which uses a sum of Gaussians to directly fit the experimental DEER data;^[13] model-free validated Tikhonov regularization using DeerAnalysis 2016;^[14] and a newer method based on the program WavPDS which uses a wavelet approach to filter out noise followed by singular value decomposition (SVD).^[15–16] The use of three distinct methods to generate $P(r)$ distributions provides a useful approach for assessing the accuracy of the resulting $P(r)$ distributions.

We first carried out DEER measurements on the PR(V82C-R1) homodimer in the absence and presence of substrates, including the RT p66/p66' precursor (Fig.2). The nitroxide label on PR is located on the side walls of the catalytic site centred around the active site residue (Asp25, here mutated to Asn) at the base and closed off at the top by two glycine-rich flaps, one from each subunit (Fig. 2A, right).^[19] The presence of V82C-R1 does not perturb the structure of PR as judged by NMR.^[20–21] There is no well-defined consensus sequence for PR cleavage, and specificity appears to depend upon the shape of the substrate envelope,^[22] as well as fit complementarity in conjunction with multiple intermolecular contacts between

the substrate, catalytic cleft and flaps to increase the lifetime of the productive complex, thereby promoting the likelihood of proteolysis.^[23] The free substrate, however, is always in the form of an intrinsically disordered, accessible region within the target protein or polyprotein and can only gain access to the catalytic cleft following transient opening of the PR laps.^[23–25]

Modelling the conformational space sampled by the V82C-R1 nitroxide label using the conformational database library MMM^[26–27] shows that in the absence of any substrate or ligand, the V82C-R1 spin label samples conformers directed both inside and outside of the catalytic cleft (Fig. 2A, middle), but in the presence of substrate the former are excluded due to steric clash with bound substrate in the active site (Fig. 2A, left). Thus the distance between pairs of V82C-R1 spin labels can be used to detect the presence of substrates/ligands/drugs within the catalytic site. The modelling is confirmed by the DEER measurements and resulting $P(r)$ distance distributions shown in Figs. 2B and C, respectively. The $P(r)$ distribution for free PR(V82C-R1) is relatively broad and centred at ~18 Å. Upon binding of two small (~500 Da) drug inhibitors, darunavir (DRV) and DMP-323 within the active site^[28] the $P(r)$ distribution narrows and shifts to ~24 Å. When the non-hydrolysable peptide substrate RPB (Ac-Thr-Ile-Nle-r-Nle-Gln-Arg-NH₂)^[18] (where r and Nle denote a reduced peptide bond and norleucine, respectively)^[18] is added, the peak of the $P(r)$ distribution is centred at ~26 Å. Only partial saturation was obtained with the hydrolysable chromogenic peptide substrate S4 (Lys-Ala-Arg-Val-Nle-(4-nitrophenylalanine)-Glu-Ala-Nle-NH₂)^[29] resulting in a narrow bound peak at ~29 Å, and a broad peak at ~20 Å corresponding to free PR.

Upon addition of the p66/p66' RT precursor to PR(V892C-R1), three peaks are observed in the $P(r)$ distribution: two narrow peaks at ~23 and ~27 Å corresponding to the bound state, and a residual broad peak at ~17 Å for the free state (Fig. 2C, lower right panel). The populations of bound and free PR from the integral of the $P(r)$ distribution are 36 and 64%, respectively. Given the total concentrations of 30 μM PR and 60 μM p66/p66' used, one can estimate that the K_D for the PR-p66/p66' complex is ~23 μM. Since peptide substrates adopt an extended conformation upon binding PR,^[24–25] the only accessible region of p66/p66' that PR can bind to is the exposed flexible linker between the connection and RNaseH domains of the open state of the p66' subunit.^[8] The presence of two bound peaks in the case of the PR-p66/p66' complex is presumably the result of two populations of nitroxide conformers compatible with the presence of the bound linker between the RNaseH and connection domains of the p66' subunit in the PR active site.

The experimental $P(r)$ distributions observed for PR(V82C-R1), both free and complexed to the various inhibitors and substrates, are not satisfactorily reproduced by the unweighted, raw ensemble of nitroxide spin label conformers generated by the program MMM^[26–27] from the structure coordinates due to multiple population subsets within the MMM-generated $P(r)$ distributions (Fig. S3). Optimizing the population weights of the MMM ensemble members, however, significantly improves agreement with the experimental $P(r)$ distributions without the need to invoke backbone shifts (see Figs. S2 and S3). Nevertheless, with the exception of free PR (V82C-R1) and possibly the complex with the S4 peptide, the calculated $P(r)$ distributions obtained by population weighting the MMM distributions

appear too broad with long components persisting and shorter ones magnified (Figs. S2 and S3). This may possibly be an artefact of the weighting algorithm and the very large number of rotamers in the raw MMM distributions. We therefore proceeded to model the nitroxide spin label conformers with the program RosettaEPR.^[30] The latter makes use of a genetic algorithm to select an optimal ensemble of rotamers, comprising anywhere between 5 and 11 members, from a crystal structure-derived rotamer library that best reproduces the experimental $P(r)$ distribution without altering the protein backbone (see SI Methods for details). The resulting RosettaEPR-generated $P(r)$ distributions are in excellent agreement with the experimental ones (Fig. S3), confirming that no changes in PR backbone conformation are required to account for the experimental $P(r)$ distributions and suggesting the existence of preferential interactions between the nitroxide spin label and drug/substrates.

In the next set of experiments, the RNase H domain of both subunits of RT p66/p66' were labeled at either G504C-R1 or Q547C-R1 (Fig. 3A). The DEER echo curves and corresponding $P(r)$ distributions are shown in Figs. 3B and C, respectively. In each instance, two distances are observed for free RT p66/p66': a shorter distance reflecting the compact state I configuration of the p66' subunit (Fig. 1A, left), and a longer distance arising from the open state II configuration (Fig. 1B, right).^[8] For G504C-R1, the peaks in the $P(r)$ distribution corresponding to states I and II of free RT p66/p66' are at ~50 and ~75 Å, respectively (Fig. 3C, top left); for Q547C-R1 these are at ~40 and ~60 Å, respectively (Fig. 3C, top right). Upon addition of PR, the G504C-R1 $P(r)$ distribution is centred about a major peak at ~58 Å, intermediate between states I and II (Fig. 3C, lower left); for Q547C-R1, the peak in the $P(r)$ distribution corresponding to state I is significantly decreased, while that corresponding to state II is increased (Fig. 3C, lower right). These data allow one to conclude that the p66' subunit in the RT p66/p66'-PR complex is largely in an open-like state, similar to state II of free p66/p66' (Fig. 3A, right).

In conclusion we have shown by DEER spectroscopy that HIV-1 PR forms a complex with the RT p66/p66' precursor. The only region of the RT p66/p66' precursor accessible to bind PR in an extended conformation is the flexible linker between the RNaseH and connection domains of the p66' subunit that is exposed in the open state II configuration,^[8] explaining why PR cleavage does not proceed further in the mature p66/p51 heterodimer.^[5, 33] The DEER data for PR (V82C-R1) indicate that the flexible linker of the p66' subunit of the p66/p66' RT precursor occupies the PR active site in a manner similar to that found with other extended peptide substrates (Fig. 2 and Figs. S2–S4), while the DEER data for p66/p66' spin-labeled at either G504C-R1 or Q547-R1 indicate that the configuration of the p66' subunit is shifted towards an open-like state in the complex (Fig. 3). These data provide an initial molecular characterization of the interaction of the RT p66/p66' precursor with HIV-1 protease. Given that both components of the complex are dimeric, and the fact that the complex is asymmetric, introduction of only a single nitroxide label per subunit for both p66/p66' and PR would result in the presence of 4 nitroxide labels within the complex giving rise to a total of 6 pairwise distance interactions (2 intramolecular and 4 intermolecular). In addition to complications arising from multispin effects,^[9] resolving that many distances would not be feasible given the width of the $P(r)$ distributions. Further characterization of the p66/p66'-PR complex by DEER would therefore require

heterologous spin labeling, for example with paramagnetic metal ion labeling of one partner and nitroxide labeling of the other, thereby permitting selective observation of metal-nitroxide, metal-metal, nitroxide-nitroxide distances.^[9, 34–35]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Drs. J. Baber and D. Garrett for technical support, and Dr. Theo Karamanos for help with implementing RosettaEPR. The SVD and WavPDS Software used in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P41GM103521. The work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases and by the Office of AIDS Research, NIH (DK029023 to GMC).

References

- [1]. Herschhorn A, Hizi A, Cell. Mol. Life Sci 2010, 67, 2717–2747. [PubMed: 20358252]
- [2]. Cabodevilla JF, Odriozola L, Santiago E, Martínez-Irujo JJ, Eur. J. Biochem 2001, 268, 1163–1172. [PubMed: 11231267]
- [3]. Sluis-Cremer N, Arion D, Abram ME, Parniak MA, Int. J. Biochem. Cell Biol 2004, 36, 1836–1847. [PubMed: 15183348]
- [4]. London RE, Viruses 2016, 8, E260. [PubMed: 27690082]
- [5]. Slack RL, Ilina TV, Xi Z, Giacobbi NS, Kawai G, Parniak MA, Sarafianos SG, Sluis Cremer N, Ishima R, Structure 2019, 27, 1581–1593 e1583. [PubMed: 31471129]
- [6]. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA, Science 1992, 256, 1783–1790. [PubMed: 1377403]
- [7]. Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, Arnold E, J. Mol. Biol 2009, 385, 693–713. [PubMed: 19022262]
- [8]. Schmidt T, Schwieters CD, Clore GM, Proc. Natl. Acad. Sci. U. S. A 2019, 116, 17809–17816. [PubMed: 31383767]
- [9]. Jeschke G, Ann. Rev. Phys. Chem 2012, 63, 419–446. [PubMed: 22404592]
- [10]. El Mkami H, Norman DG, Meth. Enzymol 2015, 564, 125–152.
- [11]. Schmidt T, Wälti MA, Baber JL, Hustedt EJ, Clore GM, Angew. Chem. Int. Ed Engl 2016, 55, 15905–15909. [PubMed: 27860003]
- [12]. Pannier M, Veit S, Godt A, Jeschke G, Spiess HW, J. Magn. Reson 2011, 213, 316–325. [PubMed: 22152351]
- [13]. Brandon S, Beth AH, Hustedt EJ, J. Magn. Reson 2012, 218, 93–104. [PubMed: 22578560]
- [14]. Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, Timmel CR, Hilger D, Jung H, App. Magn. Reson 2006, 30, 473–498.
- [15]. Srivastava M, Georgieva ER, Freed JH, J Phys Chem A 2017, 121, 2452–2465. [PubMed: 28257206]
- [16]. Srivastava M, Freed JH, J. Phys. Chem. A 2019, 123, 359–370. [PubMed: 30525624]
- [17]. Lapkouski M, Tian L, Miller JT, Le Grice SFJ, Yang W, Nat. Struct. Mol. Biol 2013, 20, 230–236. [PubMed: 23314251]
- [18]. Sayer JM, Liu F, Ishima R, Weber IT, Louis JM, J. Biol. Chem 2008, 283, 13459–13470. [PubMed: 18281688]
- [19]. Konvalinka J, Krausslich HG, Muller B, Virology 2015, 479–480, 403–417.
- [20]. Tang C, Louis JM, Aniana A, Suh JY, Clore GM, Nature 2008, 455, 693–696. [PubMed: 18833280]

- [21]. Deshmukh L, Louis JM, Ghirlando R, Clore GM, Proc. Natl. Acad. Sci. U. S. A 2016, 113, 12456–12461. [PubMed: 27791180]
- [22]. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA, Structure 2002, 10, 369–381. [PubMed: 12005435]
- [23]. Deshmukh L, Tugarinov V, Louis JM, Clore GM, Proc. Natl. Acad. Sci. U. S. A 2017, 114, E9855–E9862. [PubMed: 29087351]
- [24]. Louis JM, Ishima R, Torchia DA, Weber IT, Adv. Pharmacol 2007, 55, 261–298. [PubMed: 17586318]
- [25]. Nalam MN, Schiffer CA, Curr. Opin. HIV AIDS 2008, 3, 642–646. [PubMed: 19373036]
- [26]. Polyhach Y, Bordignon E, Jeschke G, Phys. Chem. Chem. Phys 2011, 13, 2356–2366. [PubMed: 21116569]
- [27]. Jeschke G, Protein Sci. 2018, 27, 76–85. [PubMed: 28799219]
- [28]. Lam PY, Jadhav PK, Eyermann CJ, Hodge CN, Ru Y, Bacheler LT, Meek JL, Otto MJ, Rayner MM, Wong YN, Science 1994, 263, 380–384. [PubMed: 8278812]
- [29]. Richards AD, Phylip LH, Farmerie WG, Scarborough PE, Alvarez A, Dunn BM, Hirel PH, Konvalinka J, Strop P, Pavlickova L, et al., J. Biol. Chem 1990, 265, 7733–7736. [PubMed: 2186027]
- [30]. Alexander NS, Stein RA, Koteiche HA, Kaufmann KW, McHaourab HS, Meiler J, PLoS One 2013, 8, e72851. [PubMed: 24039810]
- [31]. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA, J. Mol. Biol 2000, 301, 1207–1220. [PubMed: 10966816]
- [32]. Ishima R, Torchia DA, Lynch SM, Gronenborn AM, Louis JM, J. Biol. Chem 2003, 278, 43311–43319. [PubMed: 12933791]
- [33]. Konnyu B, Sadiq SK, Turanyi T, Hirmondo R, Muller B, Krausslich HG, Coveney PV, Muller V, PLoS Comput. Biol 2013, 9, e1003103. [PubMed: 23754941]
- [34]. Narr E, Godt A, Jeschke G, Angew. Chem. Int. Ed. Engl 2002, 41, 3907–3910. [PubMed: 12386888]
- [35]. Lueders P, Jeschke G, Yulikov M, J. Phys. Chem. Lett 2011, 2, 604–609.

Figure 1.

Maturation of HIV-1 RT. HIV-1 RT is initially produced as an asymmetric p66/p66' homodimer. Only the RNase H domain (blue ribbon) in the open state of the p66' subunit is exposed and accessible to proteolytic attack.^[8] Cleavage of the p66' RNase H domain generates the mature p66/p51 heterodimer. The PDB code for the coordinates of p66/p51 and PR are 4B3O^[17] and 3BVB,^[18] respectively.

Figure 2.

Binding of the PR(V82C-R1) homodimer to p66/p66' demonstrated by DEER. (A) Location of MTSL nitroxide labels (red circles, left), and raw rotamer distribution of the nitroxide MTSL label in the absence (centre) and presence (right) of a peptide substrate modelled using the program MMM.^[27] The PDB code used for PR in the absence and presence of the S4 peptide substrate is 1F7A,^[31] with the substrate coordinates removed for the former. (B) Q-band DEER echo curves (recorded at 50 K) of 30 μM PR(V82C-R1) dimer in the absence or presence of ligand (at a molar ratio of 1:2 PR dimer to inhibitor/substrate). The raw and background-corrected DEER echo curves are shown in grey and blue, respectively. The background is represented by a best-fit exponential in black. To optimize data point sampling for the relatively short distances being probed, the DEER echo curve was sampled with 237 points up to a maximum dipolar evolution time ($t_{\max}$) of 0.95 μs, with the position of the ELDOR pump incremented in steps of $\tau = 4$ ns. Total acquisition times ranged from 12–24 hrs. (C) $P(r)$ distributions for PR(V82C-R1) free and in the presence of drugs (DRV

and DMP-323), peptide substrates (RPB and S4), and RT p66/p66'. The position of the peak of the $P(r)$ distribution for free PR (V82C-R1) is indicated by the dashed red line. The binding affinity of the inhibitors/substrates to PR increases in the order $S4 < RBP < DMP < DRV$.^[18, 32] The data were analysed using the programs DD (green),^[13] Deer Analysis (DA, blue)^[14] and WavPDS/SVD (red).^[15–16] The results of DeerAnalysis^[14] validation for $P(r)$ distributions obtained by Tikhonov regularization are shown in Fig. S1. For $t_{\max} = 0.95 \mu\text{s}$, the upper limit for an accurate mean distance determination is $50(t_{\max}/2)^{1/3} = 39 \text{ \AA}$; the upper limit for an accurate determination of the width of the $P(r)$ distribution is $40(t_{\max}/2)^{1/3} \sim 31 \text{ \AA}$.^[9] The experimentally observed mean distances are all less than $30 \text{ \AA}$; and, from the structure of PR, the maximum possible separation between V81C-R1 nitroxide spin labels is less than $35 \text{ \AA}$ (Fig. S2). A comparison with the predicted $P(r)$ distributions, both raw and population weighted, obtained using the program MMM^[27] are shown in Figs. S2 and S3, and a comparison with the calculated RosettaEPR^[30] $P(r)$ distributions obtained using a genetic algorithm to select an optimal ensemble of nitroxide rotamers (comprising 5 to 11 members) from a crystal structure-derived rotamer library is shown in Fig. S4.

Figure 3.

HIV-1 PR binds preferentially to the open state of RT p66/p66'. (A) Closed (state I, left) and open (state II, right) conformations of the p66' subunit in the p66/p66' heterodimer.^[8] The location of the nitroxide labels are indicated by the lilac (G504C-R1) and light blue (Q547C-R1) circles. The RNase H domain and linker of the p66' subunit are shown as a ribbon; the complete p66 subunit including its RNase H domain and remainder of the p66' subunit (thumb, palm, finger and connection domains) are shown as molecular surfaces in red and blue, respectively. The linker and RNase H domain of p66' are only accessible to PR in the open state II configuration of the p66' subunit. (B) Q-band DEER echo curves (recorded at 50 K) of 30 μM RT p66/p66' (G504C-R1 or Q547C-R1) dimer in the absence or presence of 60 μM PR dimer. The raw and background-corrected DEER echo curves are shown in grey and blue, respectively. The background is represented by a best-fit exponential in black. (C) $P(r)$ distance distributions derived from the DEER data. The peak positions corresponding to states I and II of free RT p66/p66' are indicated by the light and dark blue dashed lines, respectively. The data were analysed using the programs DD (green),^[13] DeerAnalysis (DA, blue)^[14] and WavPDS/SVD (red).^[15–16] The maximum dipolar evolution time $t_{\max}$ for the DEER echo curves ranges from 13–14 μs, permitting accurate distances to be determined up to a limit of $50(t_{\max}/2)^{1/3} \sim 93\text{--}96 \text{ \AA}$.^[9] The results of DeerAnalysis^[14] validation for $P(r)$ distributions obtained by Tikhonov regularization are shown in Fig. S5.