Casper Frederiksen, MSc, PhD
Copenhagen HIV Programme
Tel: +45 35 45 57 74
Fax: +45 34 45 57 58
E-mail: cmf@cphiv.dk

A Bioinformatic Approach to Identify New Potential Resistance Relevant Amino Acid Substitutions (AAS) in HIV-1 Protease

CM Frederiksen[1], J Kjær[1], A Cozzi-Lepri[2], Z Fox[1,2] and JD Lundgren[1,3]

[1]Copenhagen HIV Programme, University of Copenhagen, Copenhagen, Denmark; [2]HIV Epidemiology & Biostatistics Group, Research Department of Infection & Population Health, University College London, London, UK; [3]Centre for Viral Diseases/KMA, Rigshospitalet, Copenhagen, Denmark

BACKGROUND

Predicting potential drug resistance mutations is important when evaluating protein-drug interactions of potential new antiviral drugs. HIV-1 Protease belongs to the large Retroviral Aspartyl Protease (RVP) family, we suggest using evolutionary information inherited in the family to investigate the most likely evolutionary amino acid substitutions (AAS).

METHODS

Figure 1 shows the structure of the protein complexes in the RVP family (pfam: PF00077, 54135 sequences) (Finn et al., 2006). **Figure 1A** shows the structure of one HIV-1 Protease complex. **Figure 1B** show an alignment (Chimera, Pettersen et al., 2004) of 30 randomly selected HIV-1 Protease complex structures from the RVP family. The structural alignment in **figure 2B** is similar to that of **figure 1B** and shows that the overall structure is conserved within the family. In this work we used a Hidden Markov Model (HMM) trained on sequences from the RVP family to estimate plausible PI resistant-associated AAS within the HIV-1 Protease. The HMM is a dynamic Bayesian network that can be used to model the order of amino acids in a protein sequence. The emission probability, stored in the model, gives the probability for a particular amino acid occurring at a particular position. From the HMM we extracted emission probabilities for each possible AAS at the 38 positions (**figure 2**) reported in the IAS drug resistance listing for HIV-1 Protease December 2008 version (Johnson et al., 2008), since they are known as positions harboring potential drug-resistant AAS. The HMM is based on curated and representative sequences from the RVP family (Finn et al., 2006).

RESULTS

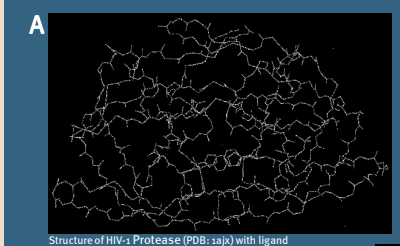
Theoretically 760 AAS (20x38) are possible for the 38 evaluated positions within the HIV-1 Protease. Of these, the RVP-HMM detected a total of 229 AAS (30.1%) with a probability above 1/20 (0.05). Of the 229 AAS, 51 (70%) were among the 73 AAS included in the IAS listing as PI-resistant mutations, leaving 178 AAS with probabilities ›0.05 as evolutionary plausible. **Figure 3** show the probability distribution of all 20 amino acids for the 38 evaluated positions. Inspecting **figure 3** show that some positions in the RVP family have high variability of AAS. One example is position 10 (L10) in **figure 3** where several amino acids have similar probabilities, and thus allow a higher number of AAS. An example of low variability is position 47 (I47) in **figure 3**. Here the amino acids I, L and V are dominant for that position. This information could be important when evaluating interactions between proteins and potential new antiviral drugs, since positions such as L10 are more likely to change and allow for increased drug-resistance.

CONCLUSION

Based on exploration of the RVP family by HMM, 70% of the established PI-resistant associated AAS could be predicted to occur. Additionally, 178 AAS were identified as evolutionary plausible and potentially could result in drug-resistance. In conclusion, we provide a probability landscape (**figure 3**) of plausible/unfavorable AAS based on inherited structure through evolution and genetic distance, which could prove useful for future drug design.

Figure 1

HIV-1 Protease Complex



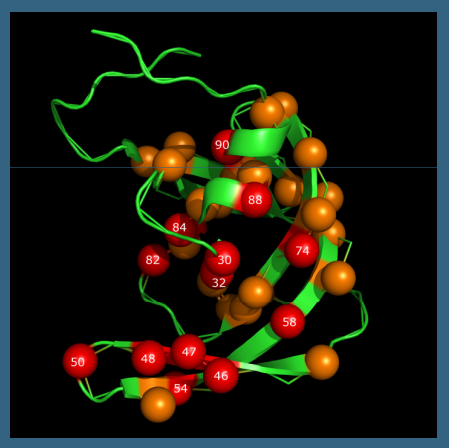
Structure of HIV-1 Protease (PDB: 1ajx) with ligand

Pettersen et al., 2004

Alignment of 30 random selected HIV-1 Protease complex structures aligned with Chimera

Figure 2

HIV-1 Protease, Chain A



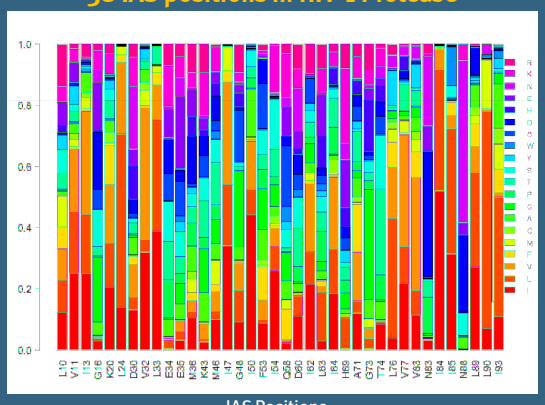
38 IAS positions

L10	Q58
V11	D60
I13	I62
G16	L63
K20	I64
L24	H69
D30	A71
V32	G73
L33	T74
E34	L76
E35	V77
M36	V83
K43	N83
M46	I84
I47	I85
G48	N88
I50	L89
F53	L90
I54	I93

The spheres show the location of 38 IAS positions (Johnson et al. 2008) in the HIV-1 Protease.
The red spheres indicate major mutations. The orange spheres indicate minor mutations.
The number on the red spheres indicate the position in the HIV-1 Protease sequence.

Figure 3

Probability Distribution for each Amino Acid at 38 IAS positions in HIV-1 Protease



Probability distribution for each amino acid extracted from the HMM-RVP family at selected IAS positions. Each color represents an amino acid. The larger the coloured bar is, the higher is the probability for the amino acid with the corresponding colour key to occur at that position.

Download poster at: www.cphiv.dk

REFERENCES

Pfam: Clans, Web Tools and Services. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL, Bateman A.
NAR, 2006, 34:D247-D25.

UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE.
J Comput Chem. 2004 Oct;25(13):1605-12.

Update of the Drug Resistance Mutations in HIV-1. Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, Pillay D, Schapiro JM, Richman DD.
Top HIV Med. 2008 Dec;16(5):138-45

Rigshospitalet

CHIP COPENHAGEN HIV PROGRAMME