European Journal of Medicinal Chemistry 54 (2012) 159–174

Contents lists available at SciVerse ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

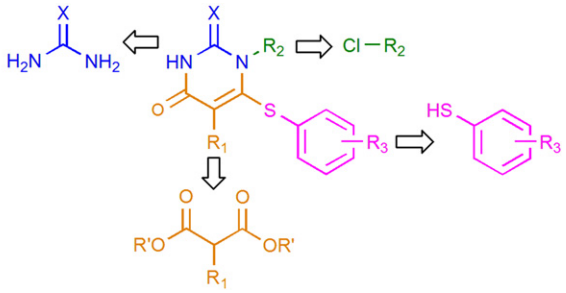
Diverse combinatorial design, synthesis and *in vitro* evaluation of new HEPT analogues as potential non-nucleoside HIV-1 reverse transcription inhibitors

Raimon Puig-de-la-Bellacasa [a], Laura Giménez [a], Sofia Pettersson [a], Rosalia Pascual [a], Encarna Gonzalo [b], José A. Esté [b], Bonaventura Clotet [b], José I. Borrell [a], Jordi Teixidó [a,*]

[a] Grup d'Enginyeria Molecular, Institut Químic de Sarrià, Universitat Ramon Llull, Via Augusta 390, Barcelona E-08017, Spain
[b] IrsiCaixa, Hospital Universitari Germans Trias i Pujol, Universitat Autònoma de Barcelona, Badalona E-08916, Spain

GRAPHICAL ABSTRACT

In a previous work, a huge virtual library of 125.396 HEPT analogues was reduced to a fully representative 560 compound combinatorial library. Now we report the synthesis and activities of a small combinatorial sub-library, confirming our assumptions that diverse combinatorial sub-library design minimizes synthetic efforts and maximizes activity range.



HIGHLIGHTS

- A library of HEPT analogues based on a rational diversity analysis is synthesized.
- The synthesized compounds are evaluated against HIV-1.
- The compounds showed a huge diversity of inhibitory activities against HIV-1.
- Some compounds showed high activity and selectivity index values.
- Diverse combionatorial sub-library design minimizes synthesis and maximizes activity.

ARTICLE INFO

Article history:
Received 1 December 2011
Received in revised form
26 April 2012
Accepted 26 April 2012
Available online 4 May 2012

ABSTRACT

New analogues of 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT) were synthesized and evaluated for their *in vitro* activities against HIV-1 in MT-4 cell cultures. Chemical diversity was introduced in 4 of the six positions of the core and the influence of each substituent was studied. This library was built on the basis of a rational diversity analysis with the objective of maximizing diversity and thus, the activity range with a minimum number of synthesized compounds. Among them, **2**{*1,2,3,1*} and **2**{*1,2,3,4*} exhibited the most potent anti-HIV-1 activities (EC_{50} = 0.015 μg/mL; 0.046 μM, SI >1667) and

Abbreviations: HEPT, 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine; CC, combinatorial chemistry; HIV, human immunodeficiency virus; HIV-1, human immunodeficiency virus type-1; NNRTI, non-nucleside reverse transcriptase inhibitor; PCA, principal component analysis; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; SEM, 2-trimethylsilylethoxymethyl; BTAC, benzyltriethylammonium chloride; BSA, *N,O*-bis(trimethylsilyl)acetamide; TBAI, tetrabutylammonium iodide; γ-PCC, picolinium chlorochromate; LR, Lawesson's reagent; NVP, nevirapine; AZT, azidothymidine; EFV, efavirenz.

* Corresponding author. Tel.: +34 932672000; fax: +34 932056265.
E-mail addresses: raimonpc@gmail.com (R. Puig-de-la-Bellacasa), laura.gimenez@ono.com (L. Giménez), ps.sofia@gmail.com (S. Pettersson), rpascual@esteve.es (R. Pascual), egonzalo@irsicaixa.es (E. Gonzalo), jaeste@irsicaixa.es (J.A. Esté), bclotet@irsicaixa.es (B. Clotet), j.i.borrell@iqs.url.edu (J.I. Borrell), jordi.teixido@iqs.url.edu (J. Teixidó).

0223-5234/$ – see front matter
doi:10.1016/j.ejmech.2012.04.038

Keywords:
Combinatorial design
Diversity selection HIV-1
NNRTIs HEPT
Antiviral activity

($EC_{50} = 0.025$ μg/mL; 0.086 μM, SI >1000), respectively, which were about 71-fold and 38-fold more active than the reference compound HEPT ($EC_{50} = 1.01$ μg/mL; 3.27 μM, SI >25).

1. Introduction

Combinatorial Chemistry (CC) has been one of the major advances introduced in medicinal chemistry during the two past decades. The concept of diversity drove the construction of large combinatorial libraries around almost any possible scaffold. The use of parallel synthesis allowed the synthesis of massive collections of compounds, even as mixtures, that were tested for relevant biological activities. However, the results obtained were worse than expected with a large number of false positives. Consequently, the use of computational selection methodologies prior to synthesis has been included as a mandatory step in CC [1,2]. In this context, our group developed Pralins (Program for Rational Analysis of Libraries In Silico) [3], a program aimed to design smaller libraries but covering the same chemical space than the complete ones. The present paper deals with its application in the field of HEPT analogues.

1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT) (**1**, Fig. 1) was discovered in 1989 by Tanaka et al. [4,5] being the first non-nucleoside reverse transcriptase inhibitor (NNRTI) ever described that was specifically targeted at the HIV-1 Reverse Transcriptase (RT) (the enzyme responsible for catalyzing the synthesis of double strain viral DNA and therefore for the replication of the virus) at an allosteric binding site [6]. Due to the high rate of virus mutation, it is important to research new molecules with improved activity and minor vulnerability to resistance [7]. A large collection of HEPT analogues **2** has been reported [5,8–13] where new substituents were chosen by medicinal chemistry intuition and synthetic accessibility.

In a previous work, we explored a large combinatorial virtual library of 125,396 HEPT analogues **2** built, following a Markush formula [14], by combining all fragments present in the published 180-compound HEPT family. There we analyzed that randomly picking 180 compounds from the library let to better coverage than the 180 synthesized compounds and that a careful selection of only 24 compounds should represent the 90% of the population. We applied Pralins to select combinatorial sub-libraries with four diversity centers (substitution at C-2 and C-5, the side chain linked to position N1 and the thiophenol ring linked to C6) showing that optimal combinatorial sub-libraries selection ranging from 24 up to $2 \times 7 \times 5 \times 8 = 560$ compounds should span from the 90% to the entirely chemical space [15].

With that objective in mind, our group proposed a single synthetic scheme (Scheme 1) for the combinatorial synthesis of HEPT analogues. Now we wish to report the results obtained in the synthesis and anti-HIV-1 evaluation of a sub-library of these new analogues and their diversity space coverage assessed in terms of broad biological activity.

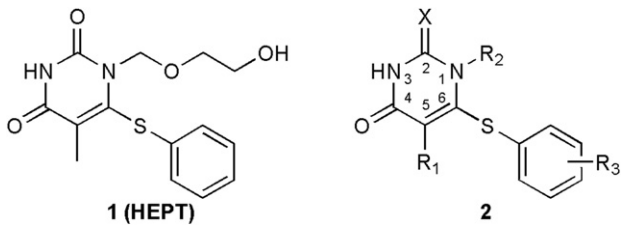


Fig. 1. HEPT and general structure of analogues **2**.

2. Results and discussion

2.1. Design and selection of the combinatorial sub-library

In order to select the sub-library to be synthesized we used our standard protocol which include the following steps.

First, we construct the complete combinatorial library by using all the relevant available fragments (typically coming from commercial reagents or from previously designed molecules). As stated before, in this case we used all fragments present in the published 180-compound HEPT family. The library enumeration, conversion of the models to 3D and geometry optimization is carried out by using either Molecular Mechanics protocols (i.e. MOE [16], Discovery Studio [17], etc.) or rule-based systems such as Corina [18,19].

Second, the chemical space is described by calculating the desired molecular descriptors by a convenient computational software suite (Discovery Studio [17], MOE [16], Dragon [20], ADRIANA.Code [21], etc.). In this case we used 9 standard combinatorial chemistry descriptors (area, molecular volume, molecular weight, radius of gyration, density, principal moment of inertia, number of rotable bonds, number of hydrogen-bond acceptors and number of hydrogen-bond donors) and 96 autocorrelogram vectors provided by ADRIANA.Code [21] (formerly PETRA [22]).

Third, the dimensionality is reduced by principal component analysis (PCA) down to a number of components that account for at least 80% of the variance. In this case 12 components were enough to represent 90% of the variance.

Fourth, we divide the PCA space by an Optimum Binning algorithm implemented in Pralins until the number of occupied cells approaches the desired selection size (typically the square-root of the total number of compounds present in the library). In this case we decided to select combinatorial sub-libraries up to $2 \times 7 \times 5 \times 8 = 560$ compounds.

Fifth, we run Pralins to carry out the selections to choose combinatorial sub-libraries which represent the original library, ideally by selecting one member of each occupied cell, in the practice more than 80% of the total number of cells or 90% of the population. By using this approach, we usually found more than one combinatorial sub-library that fulfills these criteria, so a final selection can be carried out according to complementary factors such as real availability of the reagents, cost, ADMET-Lipinski rules [23], etc.

2.2. Building blocks

In this case, with the aim of synthesizing all the selected compounds by using the synthetic pathway depicted in Scheme 1, the initial 560 member sub-library was finally reduced to 64 analogues **2** ($2X \times 2R_1 \times 4R_2 \times 4R_3$) mimicking the aforementioned protocol by selecting the corresponding building blocks (see Table 1) on the basis of commercial availability, cost and compatibility with the synthetic route.

Such synthetic route was proposed as a combination of several literature protocols. Thus, in the first step, barbiturate structures **5**