Bioorganic & Medicinal Chemistry 18 (2010) 8388–8395

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Cyclic peptide inhibitors of HIV-1 integrase derived from the LEDGF/p75 protein

Zvi Hayouka [a], Mattan Hurevich [a], Aviad Levin [b], Hadar Benyamini [a], Anat Iosub [a], Michal Maes [a], Deborah E. Shalev [c], Abraham Loyter [b], Chaim Gilon [a], Assaf Friedler [a,*]

[a] Institute of Chemistry, The Hebrew University of Jerusalem, Safra Campus, Givat Ram, Jerusalem 91904, Israel
[b] Department of Biological Chemistry, The Hebrew University of Jerusalem, Safra Campus, Givat Ram, Jerusalem 91904, Israel
[c] The Wolfson Centre for Applied Structural Biology, The Hebrew University of Jerusalem, Safra Campus, Givat Ram, Jerusalem 91904, Israel

ARTICLE INFO

Article history:
Received 22 July 2010
Revised 17 September 2010
Accepted 18 September 2010
Available online 1 October 2010

Keywords:
Cyclic peptides
Protein–protein interactions
Protein mimetics
NMR
HIV-1
Intergase inhibitors

ABSTRACT

Restricting linear peptides to their bioactive conformation is an attractive way of improving their stability and activity. We used a cyclic peptide library with conformational diversity for selecting an active and stable peptide that mimics the structure and activity of the HIV-1 integrase (IN) binding loop from its cellular cofactor LEDGF/p75 (residues 361–370). All peptides in the library had the same primary sequence, and differed only in their conformation. Library screening revealed that the ring size and linker structure had a huge effect on the conformation, binding and activity of the peptides. One of the cyclic peptides, *c*(MZ 4-1), was a potent and stable inhibitor of IN activity in vitro and in cells even after 8 days. The NMR structure of *c*(MZ 4-1) showed that it obtains a bioactive conformation that is similar to the parent site in LEDGF/p75.

1. Introduction

Proteinomimetics are molecules that mimic the structure and/or activity of an active site of a larger parent protein. Peptide-based molecules are often utilized to mimic native protein functions.[1] Developing structural proteinomimetics that also retain the function of the parent protein is challenging, but several such molecules have been reported.[2–9] Since inhibiting protein–protein interactions by small molecules is still a challenge,[10,11] proteinomimetics offer a promising alternative for targeting protein–protein interactions. Conformational constraints are required to convert a linear peptide into a proteinomimetic. Cyclic peptides are excellent candidates for proteinomimetics as they are relatively small and conformationally constrained.[1,12–15] Here we describe the development of cyclic peptides that mimic the Human immunodeficiency virus type 1 (HIV-1) integrase (IN)-binding loops from its cellular partner protein LEDGF/p75.

IN is an emerging target for anti-HIV drugs. It catalyzes the integration of viral DNA into the host genome, which is a crucial step in the HIV-1 replication cycle.[16–21] The first FDA-approved IN inhibitor, Raltegravir (MK-0518),[22–28] is used for treating HIV-1 as part of combination antiretroviral therapy. We recently developed five peptide-based IN inhibitors using either rational design, based on the interactions of IN with its partner proteins LEDGF/p75[29] and HIV-1 Rev,[29–31] or by combinatorial library screening.[32,33] All these peptides bound IN, inhibited its DNA binding and enzymatic activity in vitro, stabilized the IN tetramer, penetrated cells and inhibited HIV-1 replication in cells. The LEDGF 361–370 peptide, which was derived from one of the IN-binding loops of LEDGF/p75,[34] was particularly potent and inhibited HIV-1 replication in a mouse model.[35] LEDGF 361–370 also blocked the LEDGF/p75-IN interaction in infected cells by co-immunoprecipitation (Co-IP).[36] An independent study showed that LEDGF 361–370 competes with the full length LEDGF/p75 for IN binding with K_i = 4.6 μM.[37] Thus, LEDGF 361–370 was selected as a lead compound for further studies.

Linear peptides, such as LEDGF 361–370, are susceptible to rapid proteolytic degradation and thus are of limited use as therapeutic agents.[38,39] Restricting the peptide to its bioactive conformation is an attractive way of improving its stability, selectivity, activity and bioavailability.[1,40–42] This can be achieved using various chemical modifications, such as peptide cyclization[43] and N-alkylation.[38,42,44] Backbone cyclization (BC) is a particularly useful method to enhance the stability, bioavailability and selectivity of peptides, since it combines both N-alkylation and cyclization.[1,2,14,15,40,41,45–48] BC conformational peptide libraries, in which all peptides have exactly the same amino acid sequence,

Abbreviations: AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus; IN, integrase; LEDGF, lens epithelium-derived growth factor; IBD, integrase binding domain; BC, backbone cyclic; MOI, multiplicity of infection; BTC, bis(trichloromethyl) carbonate; RMSD, root mean square deviation; TDW, triple distilled water.

* Corresponding author. Tel.: +972 2 6585746; fax: +972 2 6585345.
E-mail address: assaf@chem.ch.huji.ac.il (A. Friedler).

0968-0896/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2010.09.046

are used to identify the cyclization mode that will confer a bioactive conformation on the peptide.[12,40,41,45,49,50] Conformational diversity is achieved by varying the position, size and chemical composition of the bridge. Screening BC peptide libraries resulted in stable cyclic peptides that mimic the structure of the parent protein and are able to inhibit protein–protein interactions[12–14] or inhibit protein activity.[15] Here, screening a BC conformational library of LEDGF 361–370 derivatives resulted in a peptide, termed *c*(MZ 4-1), which bound IN with an affinity in the low micromolar range, penetrated cells and inhibited IN catalytic activity in vitro and HIV-1 replication in cells with the same potency but with significantly improved stability compared to the linear parent peptide LEDGF 361–370. The solution NMR structures of *c*(MZ 4-1), solved both in its free form and when in interaction with IN, revealed that its bioactive conformation resembles the conformation of the corresponding residues in LEDGF/p75. Recently the crystal structure of the Prototype Foamy Virus (PFV)-IN tetramer, which is homologous to the HIV-1 IN, was solved in the presence and absence of the two potent IN inhibitors MK-0518 and GS-9137.[51] Our structural alignment results showed that *c*(MZ 4-1) has a different binding site on IN compared to these two IN inhibitors. Our results demonstrate the conversion of a rationally designed unstructured linear peptide to a cyclic peptide that is able to mimic a bioactive conformation. We conclude that *c*(MZ 4-1) is a stable and potent IN inhibitor that mimics the conformation of the parent sequence in the LEDGF protein and targets a site in IN that is different from the binding site for the IN inhibitors currently in clinical use. The methodology used here can be generally applied for developing stable peptide inhibitors of protein–protein interactions

2. Materials and methods

2.1. Cyclic peptides library synthesis, labeling, and purification

The cyclic peptides were synthesized on a Liberty Microwave-Assisted Peptide Synthesizer (CEM) using standard Fmoc chemistry. *N*-α-Fmoc-L-aspartic acid α-allyl ester or *N*-α-Fmoc-L-glutamic acid α-allyl ester were added to the parent LEDGF 361–370 sequence at its N-terminus. Protected non-natural glycine derivatives bearing alkyl chains of various lengths attached to the alpha nitrogen (Alloc-Glycine Building Unit, AGBU) were incorporated at the C-termini of peptides of the BC peptide library.[52,53] Alloc and allyl were selectively removed manually. Peptides were cyclized manually by stirring the resin with the coupling reagent 6-chloro-benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphoniumhexafluorophosphate (Pyclock) (7 equiv) (Luxembourg bio technologies, Israel) and DIPEA (14 equiv) in *N*-methyl-2-pyrrolidone overnight. The peptides were purified on a Merck Hitachi HPLC using a reverse-phase C8 semi-preparative column (Vydac) with a gradient from 5% to 60% acetonitrile in water (both containing 0.001% (v/v) trifluoroacetic acid) and ana-

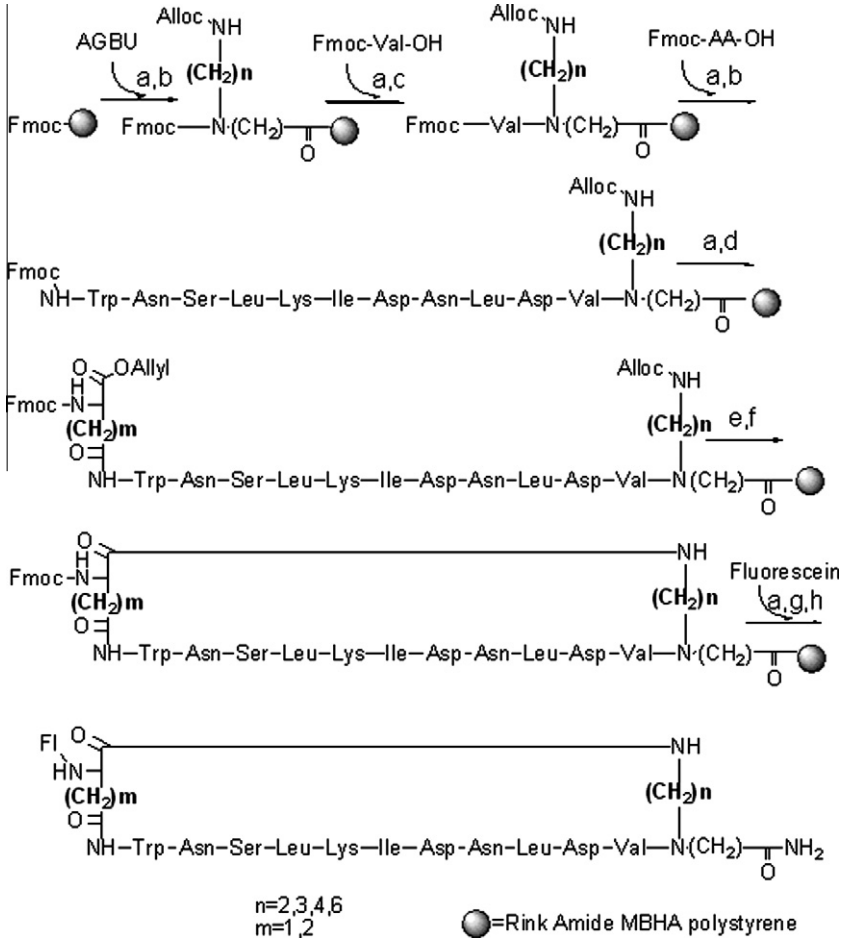

Figure 1. Design and synthesis of the BC LEDGF 361–370-derived peptide library. The peptides were synthesized using a microwave-assisted peptide synthesizer except for steps e–h that were performed manually. *n* and *m* represent the number of methylene groups in the bridge: *n* = 2, 3, 4, 6 and *m* = 1, 2. For example, *c*(MZ 4-1) represents a cyclic peptide in which *n* = 4 and *m* = 1. Reagents and conditions: (a) 20% piperidine/NMP; (b) Fmoc-AA-OH, HBTU, DIPEA; (c) Fmoc-Val-OH, HATU, DIPEA; (d) *N*-α-Fmoc-L-aspartic acid α-allyl ester or *N*-α-Fmoc-L-glutamic acid α-allyl ester, HBTU, DIPEA; (e) $(PPh_3)_4Pd(0)$, $PhSiH_3$; (f) Pyclock, DIPEA; (g) fluorescein, HBTU, DIPEA; (h) TFA, TIPS, TDW.

Table 1
LEDGF-derived cyclic peptides binding to IN and their effect on IN catalytic activity

Peptide	n[a]	m[a]	Binding affinity to IN[b] (μM)	Hill coefficient	IN inhibition[c] (%)
LEDGF 361–370			5.1 ± 0.1	3.6 ± 0.3	83
c(MZ 2-1)	2	1	5.2 ± 0.3	2.2 ± 0.2	51
c(MZ 2-2)	2	2	6.3 ± 0.5	1.6 ± 0.4	54
c(MZ3-1)	3	1	6.5 ± 0.6	1.8 ± 0.1	44
c(MZ 3-2)	3	2	4.0 ± 0.1	3.9 ± 0.9	19
c(MZ 4-1)	4	1	5.0 ± 0.4	3.4 ± 0.9	90
c(MZ 4-2)	4	2	3.1 ± 0.3	3.6 ± 0.5	65
c(MZ 6-1)	6	1	3.3 ± 0.4	3.8 ± 0.8	48
c(MZ 6-2)	6	2	3.8 ± 0.5	3.1 ± 0.6	48

[a] n and m refer to the length of the alkyl chains, as described in Figure 1.
[b] The binding affinities were determined by fluorescence anisotropy using the BC fluorescein-labeled peptides. IN (60 μM) was titrated into the fluorescein-labeled BC peptides (100 nM). Data were fit to the Hill equation (Fig. 2A).
[c] Values are taken from the data in Figure 2B.

lyzed using MALDI-TOF MS. Peptide concentrations were determined using a UV spectrophotometer (Shimadzu) as described[29] (see Fig. 1 and Table 1).

2.2. Protein expression and purification

Two types of full-length histidine-tagged IN were used: (i) IN that contains two point mutations (F185K and C280S) that moderately increase its solubility while preserving its in vitro catalytic activity.[54] This IN was used for the in vitro integration assay; (ii) For the binding studies we used a histidine-tagged IN with five mutations (C56S, W131D, F139D, F185K, and C280S) that increase its solubility, which was expressed and purified essentially as previously described.[55]

2.3. Fluorescence anisotropy

Measurements were performed at 10 °C using a Perkin–Elmer LS-55 luminescence spectrofluorimeter equipped with a Hamilton Microlab 500 dispenser.[56,57] The fluorescein-labeled LEDGF 361–370-derived cyclic peptide (1 ml, 100 nM in 20 mM Tris buffer pH 7.4, 185 mM NaCl) was placed in a cuvette, and the non-labeled IN protein (200 μL, ~50 μM) was titrated into it in 20 aliquots of 10 μL at 1 min intervals. The total fluorescence and anisotropy were measured after each addition using an excitation wavelength of 480 nm and an emission wavelength of 530 nm. Data were fit to the Hill equation:

$$R = R_0 + \frac{\Delta R * (K_a^n * [\mathrm{IN}]^n)}{1 + K_a^n * [\mathrm{IN}]^n}$$

where R is measured anisotropy, ΔR is the amplitude of the anisotropy change from R_0 (free peptide) to peptide in complex, [IN] is the added concentration of IN, and K_a is the association constant.[29]

2.4. Quantitative estimation of IN catalytic activity in vitro

Determination of the IN enzymatic activity by a quantitative assay was performed as described.[17,30] In brief, 20 μM of the linear LEDGF 361–370 parent peptide (IN-peptide molar ratio 1:50) inhibited the IN catalytic activity by ~85%. This molar ratio was used for all experiments with the cyclic peptides.

2.5. NMR measurements

A 1 mM solution of cyclic peptide in 20 mM phosphate buffer, 100 mM sodium chloride, 20.1 mM glycerol with 10% v/v D_2O was prepared from the lyophilized powder. The solution had an apparent pH of 6.87 (for further details see Supplementary data).

2.6. Docking

The PatchDock algorithm[58,59] was used to suggest models for the interaction between *c*(MZ 4-1) and the dimer of HIV-IN. Coordinates obtained by NMR in the current study were used for *c*(MZ 4-1). For the cyclic peptide, docking models with interactions of residues 366–369 were selected based on the NMR observations from the current study. For structure alignment we used MultiProt.[60]

2.7. Trypsin stability assay

The trypsin stability assay was performed as described in Ref.[61] Four hundred microliters of each peptide (1 mM) dissolved in 200 mM NH_4HCO_3 buffer solution (pH 8) were mixed with 1 μL of trypsin and chymotrypsin (porcine pancreas) solution (2.5 mg/ml). The indicted peptides were incubated at 37 °C, 30 μL samples were taken every 30 min and mixed with 30 μL of 2% TFA and 30% ACN in water. Samples were analyzed by MALDI-TOF MS.

2.8. Cell penetration experiments

The fluorescein-labeled peptides (10 μM in PBS) were incubated with HeLa cells for 2 h at 37 °C. After three washes in PBS, the cells were visualized by confocal microscopy.[29]

2.9. HIV-1 titration by multinuclear activation of a galactosidase indicator (MAGI) assay

Quantitative titration of HIV-1 was carried out using the MAGI assay with the indicated peptides at several concentrations as shown at Figure 3B, as described.[62]

2.10. Quantitative estimation of HIV-1 infection by determination of extracellular p24 lymphoid cells

Cells were incubated with the indicated peptides (12.5 μM) for 2 h and following infection with wild-type HIV-1 at a MOI of 0.01, the cells were incubated for 8 days, or 48 h at a MOI of 1 (single dose treatment). The amount of p24 protein was estimated in the cell medium every 2 days exactly as described.[30] For the multiple dose treatment a fresh dose of peptide was added to the medium every 2 days.

3. Results

3.1. Design and synthesis of a cyclic peptide library derived from LEDGF 361–370

A BC peptide library derived from the sequence of LEDGF 361–370 was designed and synthesized to select a proteinomimetic that

will adopt the bioactive conformation and possess improved stability compared to LEDGF 361–370 (Fig. 1A). For synthesizing the library, the known synthetic procedures for preparing BC peptides[12,53] were improved as follows: (1) The synthesis of BC peptides usually uses BTC as a coupling reagent, elevated temperatures and long reaction times.[47] This combination may increase racemization and lower the overall purity of the crude product. To avoid these drawbacks, we used microwave-assisted automated peptide synthesis and mild reaction conditions. Only the final steps of the synthesis, such as removing the Alloc protecting group and the cyclization, were performed manually. (2) Usually, dicarboxylic acids with varying numbers of methylene groups are used as linkers at the N-terminus of BC peptides.[52] However, using such linkers blocks the N-terminus and hence prevents the attachment of labels such as fluorescein or biotin. This precludes evaluating BC peptides using quantitative biophysical methods and studying their cell penetration. To overcome this limitation, the dicarboxylic acids were replaced by *N*-α-Fmoc-L-aspartic acid α-allyl ester or *N*-α-Fmoc-L-glutamic acid α-allyl ester, which were attached to the N-termini of the peptides. This allows both cyclization through the α-carboxyl group on the side chain and labeling the cyclic peptides by attaching fluorescein or other reagents to the free amino group on the bridge (Fig. 1). We used special building blocks termed Alloc-Glycine Building Units (AGBU).[53] Eight BC peptides with various ring sizes were synthesized and labeled with fluorescein (Fig. 1, Table 1).

3.2. The cyclic peptides bind IN with low micromolar affinity

IN binding to fluorescein-labeled cyclic peptides was characterized using fluorescence anisotropy. All peptides bound IN with affinities in the low micromolar range. The peptides with the larger ring sizes bound to the IN tetramer, with Hill coefficients around 4, as the linear LEDGF 361–370 parent peptide.[29] The BC peptides with the smaller ring sizes, Fl-*c*(MZ 2-1), Fl-*c*(MZ 2-2) and Fl-*c*(MZ 3-1), had smaller Hill coefficients of around 2 (Fig. 2A, Table 1), indicating binding to the IN dimer. The effect of the cyclic peptides on IN oligomerization equilibrium was studied using gel filtration. Free IN (10 μM) eluted at 13.5 ml, corresponding to the size of an IN dimer. In the presence of LEDGF 361–370 or *c*(MZ 4-1), the elution volume was 13.2 ml, indicating a shift towards a dimer–tetramer mixture consistent with our previous results.[29] In the presence of *c*(MZ 3-1) IN eluted as a dimer, like the free IN (Table 2). The gel filtration results are in agreement with the measured Hill coefficients.

3.3. The cyclic peptide *c*(MZ 4-1) inhibits IN catalytic activity in vitro

A quantitative enzymatic integration assay was performed to determine the effect of the BC peptides on IN catalytic activity in vitro. *c*(MZ 4-1) was the most potent inhibitor of IN in vitro, with an activity similar to linear LEDGF 361–370 (Fig. 2B). *c*(MZ 3-2), which has the same ring size as *c*(MZ 4-1) but different bridge chemistry, did not inhibit IN catalytic activity at all. This indicates that inhibition depends on the specific BC peptide conformation. Based on these findings, *c*(MZ 4-1) was selected as a lead peptide for further studies in cells.

3.4. Solution structure of the bioactive conformation of *c*(MZ 4-1)

We solved the solution structures of *c*(MZ 4-1) in the presence and absence of IN by NMR. The backbone RMSD for the final ensemble of structures of the unbound *c*(MZ 4-1) was 1.74 Å, and the lowest local RMSD for residues 366–369, which are known to

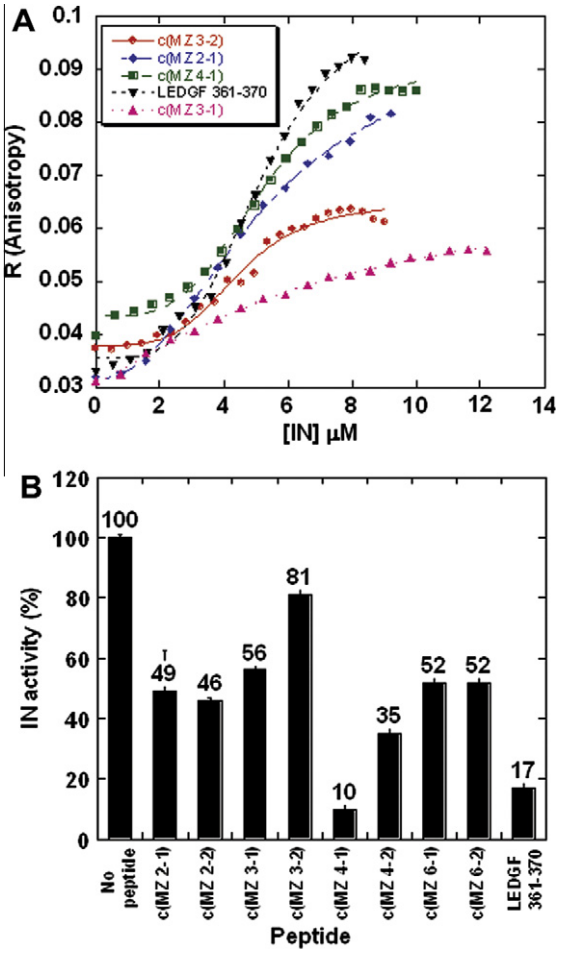

Figure 2. *c*(MZ 4-1) binds and inhibits IN catalytic activity in vitro. (a) The binding of the cyclic peptides to IN was studied using fluorescence anisotropy. IN (60 μM) was titrated into the indicated fluorescein-labeled cyclic peptide (100 nM). The binding curves were fit to the Hill equation. Binding affinities and Hill coefficients are shown in Table 1; (b) Inhibition of IN catalytic activity: IN (390 nM) was incubated in the presence of LEDGF 361–370 and its derived cyclic peptides at 1:50 IN:peptide molar ratios and the overall integration process was monitored using the quantitative assay system.

mediate IN binding of the LEDGF IBD,[34] was 0.46 Å. In the presence of IN, the backbone RMSD of the IN-bound *c*(MZ 4-1) was reduced to 1.53 Å (Fig. 3A, see also Supplementary data), while the RMSD of residues 366–369 was decreased to 0.04 Å, indicating structural stabilization of this region upon IN binding. Changes in chemical shift of *c*(MZ 4-1) upon adding IN showed larger deviations for residues 366–369 indicating that these residues participate in IN binding. The lowest energy NMR conformation of *c*(MZ 4-1) was docked to the solved structure of HIV-IN in complex with LEDGF IBD (PDB code 2b4j).[34] According to the docking model, the cyclic peptide *c*(MZ 4-1) binds IN in a similar location as the parent LEDGF 361–370 sequence[34] (Fig. 3B). The following residues of the cyclic peptide interact with HIV-IN: Asp366 was within interaction distance of Lys127 (IN chain B), supporting an electrostatic contribution. Leu368 was within interaction distance of IN chain B hydrophobic residues Gly94, Ala98, Val126, Ala128 and Ala129. Asp369 was within interaction distance of IN His171 chain B. The docking model suggests mimicry of the parent site in the LEDGF IBD with a slight shift in the location of the peptide in the IN binding pocket compared to the parent LEDGF 361–370 sequence in the protein: The residue interactions of Leu368 and Asp369 in the cyclic peptide correspond to the hydrophobic and electrostatic interactions of Ile365 and Asp366 of the full length LEDGF IBD

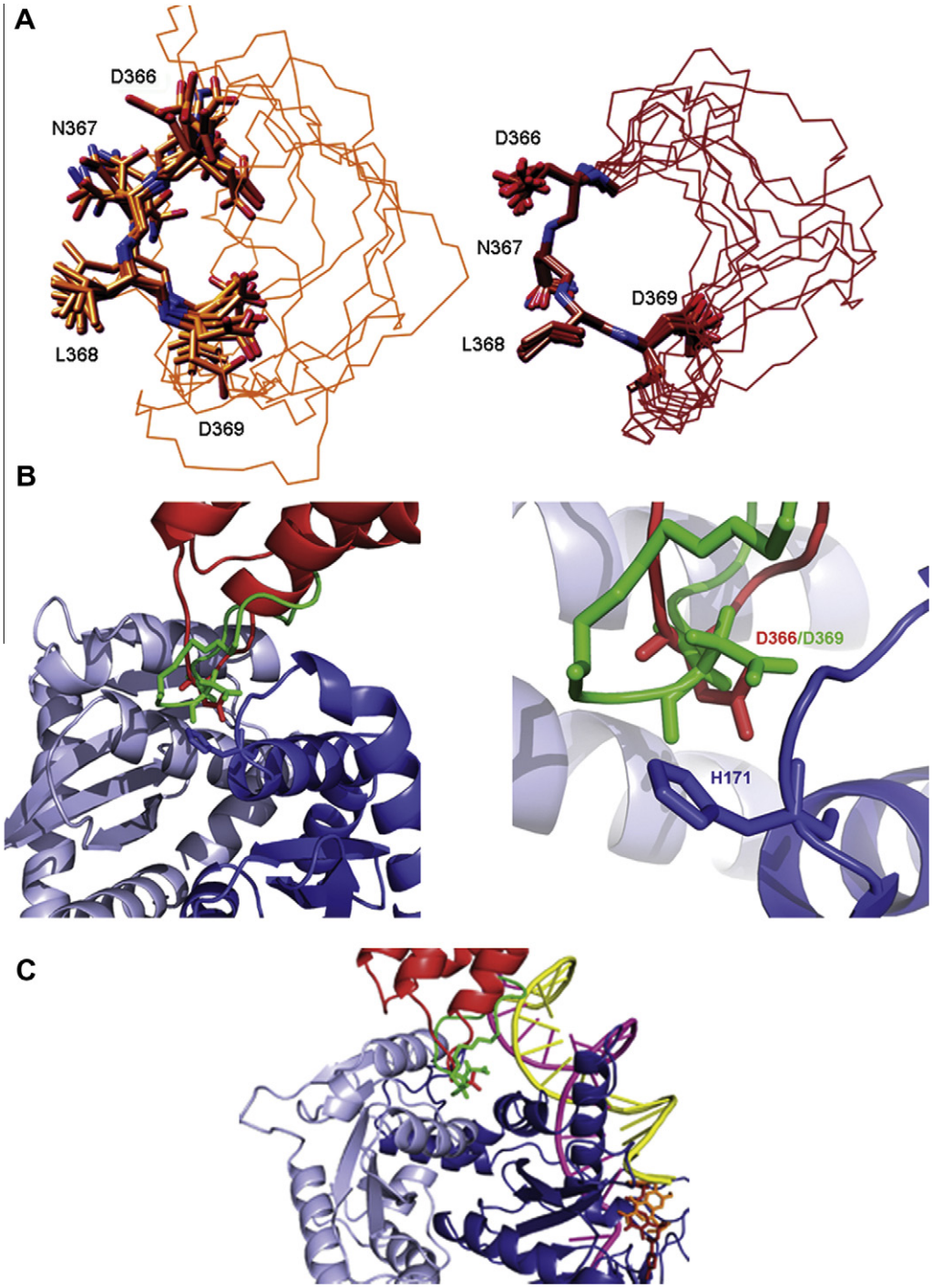

Figure 3. *c*(MZ 4-1) mimics the bioactive conformation of LEDGF/p75: NMR and docking studies. (a) The NMR solution structure ensemble of *c*(MZ 4-1) in the presence and absence of IN. (Orange, left panel) The backbone RMSD for the final ensemble of structures of the unbound *c*(MZ 4-1) was 1.74 Å, and the local RMSD for residues 366–369 (sticks) was 0.46 Å. (Red, right panel) In the presence of IN the local RMSD of the same region decreased to 0.04 Å (sticks) while the rest the cyclic peptide remains flexible. Molecular graphics were prepared using Chimera.[65] (b) *c*(MZ 4-1) mimics the LEDGF IBD binding site: a docking study. (Left panel) The active cyclic peptide *c*(MZ 4-1) was docked to the solved structure of IN core domain dimer (blue and light-blue each monomer) in complex with LEDGF/p75 IBD (red) (PDB code: 2b4j). The docking model places the cyclic peptide (green) in a similar yet slightly different location relative to the location of the parent LEDGF 361–370 in the crystal structure. (Right panel) Enlarged interaction site: D366 of LEDGF/p75 IBD (red) structurally corresponds to D369 of *c*(MZ 4-1) (green). *c*(MZ 4-1) mimicked the binding region of LEDGF/p75 IBD with a slight shift in the residues that participate in binding. D366, which was reported as an important residue at the protein level, is replaced by D369 in the peptide. (C) Different binding sites on IN for *c*(MZ 4-1), the DNA and the strand transfer inhibitors. The figure depicts the structural alignment of three IN structures: the complex of HIV-IN with LEDGF IBD with *c*(MZ 4-1) docked into it (PDB id: 2B4J[53]) and the structures of PFV-IN with strand transfer inhibitors MK-0518 and GS-9137 (PDB: 3L2T, 3L2U, respectively[51]). For clarity only the PVF-IN chains are shown. Chains A and B of PFV-IN are colored blue and light-blue, respectively. The inhibitors MK-0518 and GS-9137 are depicted in sticks and colored brown and orange. DNA strands are colored yellow and magenta. The LEDGF chain C (PDB: 2B4J) is colored red, and *c*(MZ 4-1) is colored green.

(Fig. 3B), similar to the linear LEDGF 361–370 peptide.[35] Altogether, the NMR and docking results demonstrate a successful structural mimicry of the IN-binding loop formed by residues 361–370 of LEDGF/p75.

Table 2
c(MZ 4-1) affects the IN oligomerization equilibrium[a]

Sample	Elution volume (ml)	Estimated oligomeric state of IN	MW (kDa)
BSA	13.5		67
Aldolase	12.5		158
IN 60 μM	12.8	Tetramer	128
IN 10 μM	13.5	Dimer	64
IN 10 μM with LEDGF 361–370 10 μM	13.2	Dimer–tetramer	64–128
IN 10 μM with *c*(MZ 4-1) 10 μM	13.2	Dimer–tetramer	64–128
IN 10 μM with *c*(MZ 3-1) 10 μM	13.5	Dimer	64

[a] Oligomerization of IN in the presence of various peptides was studied using analytical gel filtration. IN (10 μM) alone eluted as a dimer. In the presence of LEDGF 361–370 or *c*(MZ 4-1) IN eluted as a mixture of dimers and tetramers, and in the presence of *c*(MZ 3-1) the oligomeric state of IN was unchanged and IN eluted as a dimer.

We aligned the docking model of the IN—*c*(MZ 4-1) complex with the structures of with the HIV-1 IN—LEDGF/p75 IBD complex (PDB id: 2B4 J,[53]) and with the structures of PFV-IN in complex with the strand transfer inhibitors MK-0518 and GS-9137 (PDB: 3L2T, 3L2U, respectively[51]). Figure 3C demonstrates that the binding sites on IN for *c*(MZ 4-1) and for the small molecule inhibitors are different.

3.5. *c*(MZ 4-1) blocked viral gene expression and had a long term inhibitory effect on HIV-1 replication in cells

Six of the eight fluorescein-labeled BC peptides penetrated cells and two, Fl-*c*(MZ 3-1) and Fl-*c*(MZ 3-2), did not (Fig. 4A). The effect of the lead BC peptide *c*(MZ 4-1) on HIV-1 propagation was studied by the MAGI assay.[54] Both LEDGF 361–370 and *c*(MZ 4-1) significantly inhibited HIV-Tat mediated expression of the reporter gene in a concentration-dependent manner (Fig. 4B), indicating that the transcription of viral genes was inhibited. *c*(MZ 3-2), which has the same ring size as the lead *c*(MZ 4-1), was unable to penetrate cells. Thus *c*(MZ 6-1), which penetrated cells and moderately inhibited IN catalytic activity in vitro but did not inhibit viral gene expression, was selected as a negative control for further studies. The effect of *c*(MZ 4-1) versus *c*(MZ 6-1) on HIV-1 replication was evaluated by testing their ability to inhibit the release of the viral p24 protein in infected T-lymphoid cells. The time dependent inhibitory effect of the BC peptides compared to the linear LEDGF 361–370 peptide was determined. Each peptide was introduced either in a single dose at day 0 or in multiple doses: administering an additional dose every 2 days. The linear LEDGF 361–370 inhibited HIV-1 replication at 8 days post infection only upon multiple dose treatment but not upon single dose treatment. On the other hand, the cyclic *c*(MZ 4-1) blocked HIV-1 replication both in multiple dose treatment (>90% inhibition 8 days post infection) and in single dose treatment (80% inhibition 8 days post infection) (Fig. 4C), indicating a dramatic increase in the stability of the peptide compared to its linear parent peptide.

To verify that cyclization indeed increased the stability of the peptides, LEDGF 361–370 and *c*(MZ 4-1) were incubated at 37 °C in a mixture of trypsin and chymotrypsin and their degradation was monitored by mass spectrometry over time. LEDGF 361–370 was degraded after 30 min while *c*(MZ 4-1) started to degrade only after 150 min (data not shown). We conclude that *c*(MZ 4-1) is indeed more resistant to trypsin digestion than the linear parent LEDGF 361–370.

4. Discussion

Peptides have emerged as adequate molecules for mimicking protein binding sites, as well as inhibiting protein–protein interactions.[63] In this study we designed cyclic peptides to mimic the bioactive conformation of LEDGF/p75 binding site to IN. The cyclic peptides were extensively compared using biophysical, biochemical, and structural methods. One of the cyclic peptides, *c*(MZ 4-1), emerged as a stable and potent anti-HIV lead compound. Our structural alignment indicated that *c*(MZ 4-1) binds to a different site on IN compared to the known strand transfer inhibitors MK-0518 and GS-9137. Thus, *c*(MZ 4-1) suggests an alternative target on the IN structure and may have the potential to be coupled in the future with the FDA-approved IN inhibitors to reduce drug resistance.

4.1. Novel implications for cyclic peptide synthesis

We applied here for the first time a new semi-automated synthetic scheme for preparing BC peptide libraries. The pre-cyclic linear peptide, including the AGBU, was synthesized automatically in a Microwave-Assisted Peptide Synthesizer using mild conditions of HBTU/HOBT chemistry and only the final steps of Alloc deprotection and cyclization were conducted manually. An important novelty in the chemical structure of these BC peptides is the additional free amine moiety on the linker at the N-terminus of the peptide, which can be attached to additional chemical entities, including labeling agents such as biotin or fluorescent dyes as well as cell penetrating agents. Here, the BC peptides were labeled with fluorescein, to enable their characterization using biophysical methods and determination of their cellular uptake.

4.2. The peptide conformation affects multiple parameters in its activity

The cyclic peptides described in this work differ from each other in many properties, including their ability to inhibit IN activity, to penetrate cells and to shift the IN oligomerization equilibrium. Since all the cyclic peptides described here have the same parent sequence, the difference in properties is attributed to their different conformations. Minor changes in the ring size of peptides with identical sequences changed their cell penetration properties, suggesting that these are governed by the peptide conformation. We have previously shown that the linear LEDGF 361–370 stabilized the IN tetramer and shifted the IN oligomerization equilibrium towards the tetramer.[29] We termed peptides with such activity 'shiftides'.[29] Here, the gel filtration results and the Hill coefficients show that different BC peptides, bearing the same parent sequence, have a different effect on IN oligomerization. Most peptides with larger ring sizes stabilized the IN tetramer, while the peptides with smaller rings bound the IN dimer. A similar trend of dependence on the peptide length was observed in our previous study, where a 20-mer IN inhibitor bound the IN tetramer, while its shorter 10-mer derivatives bound the IN dimer.[33] These results indicate that the peptide conformations, and not only the sequence, affect the ability of shiftides to shift the oligomerization equilibrium of a protein.

The different cyclic peptides had different numbers of methylenes in the bridge, resulting in changes in the overall hydrophobic-

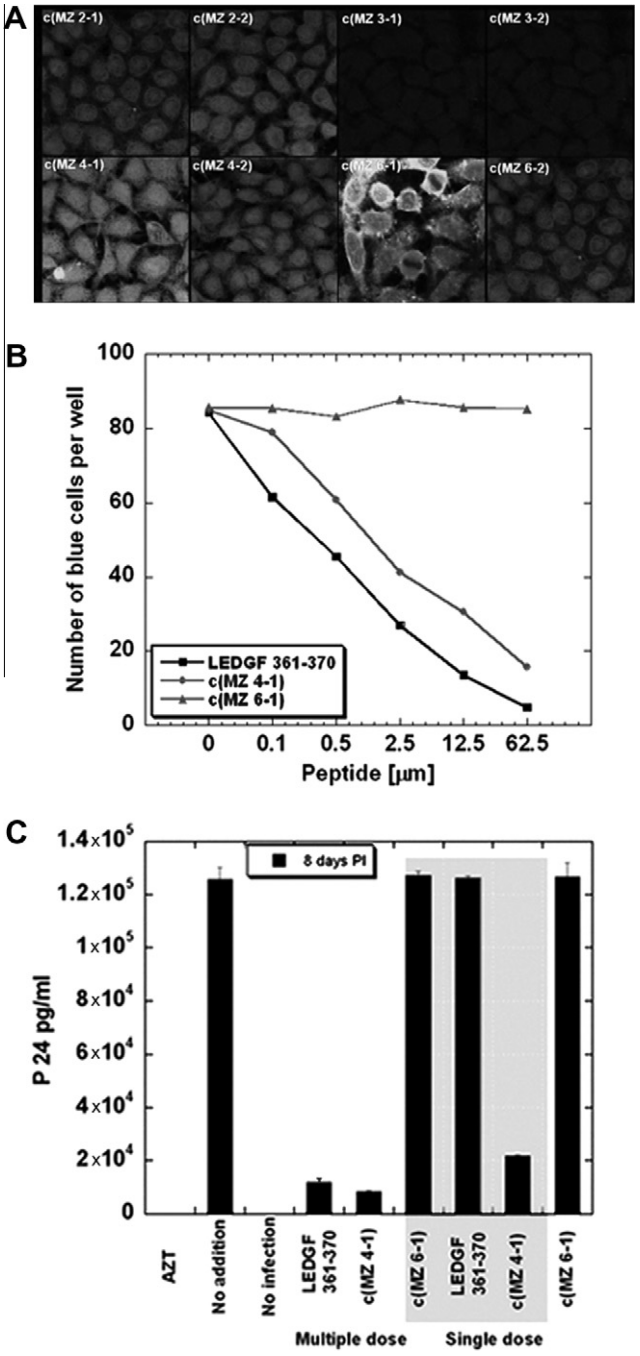

Figure 4. *c*(MZ 4-1) inhibits HIV-1 replication in cells with improved stability. (a) All BC peptides, except Fl-*c*(MZ 3-1) and Fl-*c*(MZ 3-2), penetrated cells. The peptides localized both to the cytoplasm and the nucleus, as visualized by confocal microscope. (b) *c*(MZ 4-1) inhibited TAR-mediated transcription of HIV-1 genes in TZM-bl MAGI cells in a concentration-dependent manner. (c) *c*(MZ 4-1) inhibited HIV-1 replication in cell culture after 8 days. H9 T-lymphoid cells were incubated with *c*(MZ 4-1), *c*(MZ 6-1) and LEDGF 361–370, and the total amount of the released virus was estimated based on the p24 protein content at days 2, 4, 6, and 8. Shown is the quantification of the inhibition of p24 formation in T-lymphoid cells after 8 days. To determine the time-dependent inhibitory effect of *c*(MZ 4-1), *c*(MZ 6-1), and LEDGF 361–370 the peptides were introduced either in single dose at day 0 (single dose, gray background) or in multiple dose, in which an additional dose was added every 2 days (multiple dose).

ity. However, this is unlikely to be the cause for the differential activity, since peptides with the same bridge size and hydrophobicity but different bridge chemistry showed different behavior such as *c*(MZ 4-1) and *c*(MZ 3-2).

4.3. LEDGF 361–370 derived cyclic peptides: structural and functional comparison

The NMR solution structure of the active *c*(MZ 4-1) indicated that the conformation of residues 366–369 was stabilized upon IN binding. This suggests that these residues, which are known to bind IN in the parent protein,[34] adopted a bioactive conformation in the peptide as well. Residues 366–369 in *c*(MZ 4-1) showed a low local RMSD value, describing a more confined conformational space. The observation that *c*(MZ 4-1) showed the lowest local RMSD while interacting with IN, suggests that it achieves and maintains the bioactive conformation better than the other peptides. The docking results suggest that *c*(MZ 4-1) mimics the LEDGF/p75 IBD binding site,[34] with a possible slight shift where the residue interactions of Leu368 and Asp369 in the peptide correspond to the hydrophobic and electrostatic interactions of Ile365 and Asp366 in the parent LEDGF/p75 protein, respectively. This shift has been previously shown by us for the linear LEDGF 361–370 peptide as well.[35]

Previous studies demonstrated that BC improves resistance towards enzymatic degradation.[12,49,64] The stability assays performed here, combined with the long term inhibitory effect of *c*(MZ 4-1) on HIV-1 replication in the single dose experiments, imply that *c*(MZ 4-1) is more stable than the linear parent peptide. The BC peptide *c*(MZ 4-1) retained full potency 8 days post infection and preserved the inhibitory activity, while the linear peptide degraded rapidly. These findings indicate that the cyclization stabilizes the peptide and increases its duration of activity in cells. The BC peptide is especially stabilized since it contains two modes of conformational constraint: N-methylation and cyclization. We conclude that BC is a powerful method to select the active, stable and bioactive conformation of cyclic peptides. *c*(MZ 4-1) may potentially become an anti-HIV lead compound.

5. Conclusions

In this study we demonstrated the conversion of a rationally designed unstructured linear peptide to a conformationally restricted and stable cyclic peptide that mimics the bioactive conformation. The NMR studies show a successful structural mimicry of the IN-binding loop in LEDGF/p75. Because of its ability to mimic the bioactive conformation of the IN-binding loop in LEDGF/p75, *c*(MZ 4-1) is improved relative to the linear parent peptide: It has the same activity because it mimics the bioactive conformation, and it has a significantly improved stability in vitro and in cells because of the cyclization, as shown by its ability to retain full potency 8 days post infection. The methodology shown here is general and can be applied to other protein–protein interactions.

Acknowledgments

A.F. is supported by a starting grant from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement n° 203413. A.L. is supported by the Israel Science Foundation (ISF; Grant No. 888/05). H.B is supported by a fellowship from the Israel Cancer research Foundation (ICRF). Some molecular graphics images were produced using the Chimera package from the University of California, San Francisco (supported by NIH P41 RR-01081).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2010.09.046.

References

1. Kasher, R.; Oren, D. A.; Barda, Y.; Gilon, C. *J. Mol. Biol.* **1999**, *292*, 421.
2. Murray, J. K.; Gellman, S. H. *Biopolymers* **2007**, *88*, 657.
3. Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342.
4. Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; DeGrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425.
5. Dechantsreiter, M. A.; Planker, E.; Matha, B.; Lohof, E.; Holzemann, G.; Jonczyk, A.; Goodman, S. L.; Kessler, H. *J. Med. Chem.* **1999**, *42*, 3033.
6. Wrighton, N. C.; Farrell, F. X.; Chang, R.; Kashyap, A. K.; Barbone, F. P.; Mulcahy, L. S.; Johnson, D. L.; Barrett, R. W.; Jolliffe, L. K.; Dower, W. J. *Science* **1996**, *273*, 458.
7. Laudet, B.; Barette, C.; Dulery, V.; Renaudet, O.; Dumy, P.; Metz, A.; Prudent, R.; Deshiere, A.; Dideberg, O.; Filhol, O.; Cochet, C. *Biochem. J.* **2007**, *408*, 363.
8. Franke, R.; Hirsch, T.; Eichler, J. *J. Recept. Signal Transduct. Res.* **2006**, *26*, 453.
9. Hamm, H. E.; Deretic, D.; Arendt, A.; Hargrave, P. A.; Koenig, B.; Hofmann, K. P. *Science* **1988**, *241*, 832.
10. Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Disc.* **2004**, *3*, 301.
11. Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001.
12. Hess, S.; Linde, Y.; Ovadia, O.; Safrai, E.; Shalev, D. E.; Swed, A.; Halbfinger, E.; Lapidot, T.; Winkler, I.; Gabinet, Y.; Faier, A.; Yarden, D.; Xiang, Z.; Portitlo, F. P.; Haskell-Luevano, C.; Gilon, C.; Hoffman, A. *J. Med. Chem.* **2008**, *51*, 1026.
13. Linde, Y.; Ovadia, O.; Safrai, E.; Xiang, Z.; Portillo, F. P.; Shalev, D. E.; Haskell-Luevano, C.; Hoffman, A.; Gilon, C. *Biopolymers* **2008**, *90*, 671.
14. Friedler, A.; Friedler, D.; Luedtke, N. W.; Tor, Y.; Loyter, A.; Gilon, C. *J. Biol. Chem.* **2000**, *275*, 23783.
15. Friedler, A.; Zakai, N.; Karni, O.; Broder, Y. C.; Baraz, L.; Kotler, M.; Loyter, A.; Gilon, C. *Biochemistry* **1998**, *37*, 5616.
16. Sherman, M. P.; Greene, W. C. *Microb. Infect.* **2002**, *4*, 67.
17. Engelman, A.; Mizuuchi, K.; Craigie, R. *Cell* **1991**, *67*, 1211.
18. Guiot, E.; Carayon, K.; Delelis, O.; Simon, F.; Tauc, P.; Zubin, E.; Gottikh, M.; Mouscadet, J. F.; Brochon, J. C.; Deprez, E. *J. Biol. Chem.* **2006**, *281*, 22707.
19. Deprez, E.; Tauc, P.; Leh, H.; Mouscadet, J. F.; Auclair, C.; Brochon, J. C. *Biochemistry* **2000**, *39*, 9275.
20. Deprez, E.; Tauc, P.; Leh, H.; Mouscadet, J. F.; Auclair, C.; Hawkins, M. E.; Brochon, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10090.
21. Faure, A.; Calmels, C.; Desjobert, C.; Castroviejo, M.; Caumont-Sarcos, A.; Tarrago-Litvak, L.; Litvak, S.; Parissi, V. *Nucleic Acids Res.* **2005**, *33*, 977.
22. Pommier, Y.; Johnson, A. A.; Marchand, C. *Nat. Rev. Drug Disc.* **2005**, *4*, 236.
23. Lataillade, M.; Kozal, M. J. *AIDS Patient Care STDS* **2006**, *20*, 489.
24. Craigie, R. *J. Biol. Chem.* **2001**, *276*, 23213.
25. Merschman, S. A.; Vallano, P. T.; Wenning, L. A.; Matuszewski, B. K.; Woolf, E. J. *J. Chromatogr., B. Analyt. Technol. Biomed. Life Sci.* **2007**, *857*, 15.
26. Hazuda, D. J.; Young, S. D.; Guare, J. P.; Anthony, N. J.; Gomez, R. P.; Wai, J. S.; Vacca, J. P.; Handt, L.; Motzel, S. L.; Klein, H. J.; Dornadula, G.; Danovich, R. M.; Witmer, M. V.; Wilson, K. A.; Tussey, L.; Schleif, W. A.; Gabryelski, L. S.; Jin, L.; Miller, M. D.; Casimiro, D. R.; Emini, E. A.; Shiver, J. W. *Science* **2004**, *305*, 528.
27. Markowitz, M.; Morales-Ramirez, J. O.; Nguyen, B. Y.; Kovacs, C. M.; Steigbigel, R. T.; Cooper, D. A.; Liporace, R.; Schwartz, R.; Isaacs, R.; Gilde, L. R.; Wenning, L.; Zhao, J.; Teppler, H. *J. Acquir. Immune Defic. Syndr.* **2006**, *43*, 509.
28. Evering, T. H.; Markowitz, M. *Drugs Today (Barc)* **2007**, *43*, 865.
29. Hayouka, Z.; Rosenbluh, J.; Levin, A.; Loya, S.; Lebendiker, M.; Veprintsev, D.; Kotler, M.; Hizi, A.; Loyter, A.; Friedler, A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8316.
30. Rosenbluh, J.; Hayouka, Z.; Loya, S.; Levin, A.; Armon-Omer, A.; Britan, E.; Hizi, A.; Kotler, M.; Friedler, A.; Loyter, A. *J. Biol. Chem.* **2007**, *282*, 15743.
31. Hayouka, Z.; Rosenbluh, J.; Levin, A.; Maes, M.; Loyter, A.; Friedler, A. *Biopolymers* **2008**, *90*, 481.
32. Armon-Omer, A.; Levin, A.; Hayouka, Z.; Butz, K.; Hoppe-Seyler, F.; Loya, S.; Hizi, A.; Friedler, A.; Loyter, A. *J. Mol. Biol.* **2008**, *376*, 971.
33. Maes, M.; Levin, A.; Hayouka, Z.; Shalev, D. E.; Loyter, A.; Friedler, A. *Bioorg. Med. Chem.* **2009**, *17*, 7635.
34. Cherepanov, P.; Ambrosio, A. L.; Rahman, S.; Ellenberger, T.; Engelman, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17308.
35. Hayouka, Z.; Levin, A.; Maes, M.; Hadas, E.; Shalev, D. E.; Volsky, D. J.; Loyter, A.; Friedler, A. *Biochem. Biophys. Res. Commun.* **2010**, *394*, 260.
36. Levin, A.; Rosenbluh, J.; Hayouka, Z.; Friedler, A.; Loyter, A. *Mol. Med.* **2009**.
37. Tsiang, M.; Jones, G. S.; Hung, M.; Mukund, S.; Han, B.; Liu, X.; Babaoglu, K.; Lansdon, E.; Chen, X.; Todd, J.; Cai, T.; Pagratis, N.; Sakowicz, R.; Geleziunas, R. *J. Biol. Chem.* **2009**, *284*, 33580.
38. Biron, E.; Chatterjee, J.; Ovadia, O.; Langenegger, D.; Brueggen, J.; Hoyer, D.; Schmid, H. A.; Jelinek, R.; Gilon, C.; Hoffman, A.; Kessler, H. *Angew. Chem., Int. Ed.* **2008**, *47*, 2595.
39. Woodley, J. F. *Crit. Rev. Ther. Drug Carrier Syst.* **1994**, *11*, 61.
40. Gilon, C.; Halle, D.; Chorev, M.; Selinger, Z.; Byk, G. *Biopolymers* **1991**, *31*, 745.
41. Grdadolnik, S. G.; Mierke, D. F.; Byk, G.; Zeltser, I.; Gilon, C.; Kessler, H. *J. Med. Chem.* **1994**, *37*, 2145.
42. Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. *Acc. Chem. Res.* **2008**, *41*, 1331.
43. Kessler, H. *Angew. Chem.* **1982**, *21*, 512.
44. Chatterjee, J.; Ovadia, O.; Gilon, C.; Hoffman, A.; Mierke, D.; Kessler, H. *Adv. Exp. Med. Biol.* **2009**, *611*, 109.
45. Gilon, C.; Lohof, E.; Friedler, A.; Kessler, H.; Synthesis of Cyclic Peptides, In *Houben–Weyl: Methods of Organic Chemistry, Synthesis of Peptides and Peptidomimetics*, Goodman, M.; Felix, A.; Moroder, L.; Toniolo, C., Eds.; Thieme Verlag: Stuttgart, Germany, 2003; E 22b, pp 461–542.
46. Hariton-Gazal, E.; Rosenbluh, J.; Zakai, N.; Fridkin, G.; Brack-Werner, R.; Wolff, H.; Devaux, C.; Gilon, C.; Loyter, A. *Biochemistry* **2005**, *44*, 11555.
47. Qvit, N.; Hatzubai, A.; Shalev, D. E.; Friedler, A.; Ben-Neriah, Y.; Gilon, C. *Biopolymers* **2009**, *91*, 157.
48. Baraz, L.; Friedler, A.; Blumenzweig, I.; Nussinuv, O.; Chen, N.; Steinitz, M.; Gilon, C.; Kotler, M. *FEBS Lett.* **1998**, *441*, 419.
49. Byk, G.; Halle, D.; Zeltser, I.; Bitan, G.; Selinger, Z.; Gilon, C. *J. Med. Chem.* **1996**, *39*, 3174.
50. Hariton-Gazal, E.; Friedler, D.; Friedler, A.; Zakai, N.; Gilon, C.; Loyter, A. *Biochim. Biophys. Acta* **2002**, *1594*, 234.
51. Hare, S.; Gupta, S. S.; Valkov, E.; Engelman, A.; Cherepanov, P. *Nature* **2010**, *464*, 232.
52. Barda, Y.; Cohen, N.; Lev, V.; Ben-Aroya, N.; Koch, Y.; Mishani, E.; Fridkin, M.; Gilon, C. *Nucl. Med. Biol.* **2004**, *31*, 921.
53. Hurevich, M.; Barda, Y.; Gilon, C. *Heterocycles* **2007**, *73*, 617.
54. Jenkins, T. M.; Engelman, A.; Ghirlando, R.; Craigie, R. *J. Biol. Chem.* **1996**, *271*, 7712.
55. Chen, J. C.; Krucinski, J.; Miercke, L. J.; Finer-Moore, J. S.; Tang, A. H.; Leavitt, A. D.; Stroud, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8233.
56. Friedler, A.; Hansson, L. O.; Veprintsev, D. B.; Freund, S. M.; Rippin, T. M.; Nikolova, P. V.; Proctor, M. R.; Rudiger, S.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 937.
57. Friedler, A.; Veprintsev, D. B.; Rutherford, T.; von Glos, K. I.; Fersht, A. R. *J. Biol. Chem.* **2005**, *280*, 8051.
58. Duhovny, D.; Nussinov, R.; Wolfson, H. J. Efficient Unbound Docking of Rigid Molecules In *Workshop on Algorithms in Bioinformatics*; Guigo, R., Gusfield, D., Eds.; LNCS: Rome, Italy, 2002; Vol. 2452, pp 185–200.
59. Schneidman-Duhovny, D.; Inbar, Y.; Polak, V.; Shatsky, M.; Halperin, I.; Benyamini, H.; Barzilai, A.; Dror, O.; Haspel, N.; Nussinov, R.; Wolfson, H. J. *Proteins* **2003**, *52*, 107.
60. Shatsky, M.; Nussinov, R.; Wolfson, H. J. *Proteins* **2004**, *56*, 143.
61. Tal-Gan, Y.; Freeman, N. S.; Klein, S.; Levitzki, A.; Gilon, C. *Bioorg. Med. Chem.* **2010**, *18*, 2976.
62. Levin, A.; Hayouka, Z.; Helfer, M.; Brack-Werner, R.; Friedler, A.; Loyter, A. *PLoS ONE* **2009**, *4*, e4155.
63. Eichler, J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 707.
64. Hess, S.; Ovadia, O.; Shalev, D. E.; Senderovich, H.; Qadri, B.; Yehezkel, T.; Salitra, Y.; Sheynis, T.; Jelinek, R.; Gilon, C.; Hoffman, A. *J. Med. Chem.* **2007**, *50*, 6201.
65. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.