Pharmacophore Modeling and Molecular Docking Led to the Discovery of Inhibitors of Human Immunodeficiency Virus-1 Replication Targeting the Human Cellular Aspartic Acid−Glutamic Acid−Alanine−Aspartic Acid Box Polypeptide 3

Giovanni Maga,*[,†] Federico Falchi,[‡] Anna Garbelli,[†] Amalia Belfiore,[‡] Myriam Witvrouw,[§] Fabrizio Manetti,[‡] and Maurizio Botta*[,‡]

Istituto di Genetica Molecolare, IGM-CNR, Via Abbiategrasso 207, I-27100 Pavia, Italy, Dipartimento Farmaco Chimico Tecnologico, Università degli Studi di Siena, Via Aldo Moro, I-53100 Siena, Italy, and Molecular Medicine, Katholieke Universiteit Leuven and IRC KULAK, Kapucijnenvoer 33, B-3000 Leuven, Belgium

Abstract: HIV-1 replication has been inhibited by using a compound able to target the human cellular cofactor DEAD-box ATPase DDX3, essential for HIV-1 RNA nuclear export. This compound, identified by means of a computational protocol based on pharmacophoric modeling and molecular docking calculations, represents the first small molecule with such a mechanism of action and could lay the foundations for a pioneering approach for the treatment of HIV-1 infections.

Compounds currently used for the treatment of HIV-1 infections are targeted to viral proteins.[1] However, the high intrinsic mutation and replication rates of HIV-1 led to the emergence of drug resistant strains with a consequent therapeutic failure. On this basis, cellular cofactors represent attractive new targets for HIV-1 chemotherapy, since targeting a cellular factor that is required for viral replication should help to overcome the problem of viral resistance.

The only anti-HIV-1 drug approved so far that targets a cellular protein is maraviroc,[2] an antagonist of the cellular chemokine receptor CCR5[a] that, together with CXCR4, acts as an HIV-1 coreceptor, essential for viral entry. However, maraviroc efficacy is limited to CCR5-tropic HIV-1 strains.

To allow nuclear export of its own unprocessed or partially processed RNA transcripts, HIV-1 must take control of the strict quality control pathways that mammalian cells have evolved to avoid the nuclear export of partially spliced transcripts.[3] Consequently, at this particular step of the viral life cycle, functional interactions between viral and cellular proteins likely play a pivotal role. Any drug targeting this step will act at an earlier level than protease inhibitors, shutting down viral proliferation before structural viral proteins (such as Gag-Pol and Env) or even full length genomic RNA has been produced in the infected cells. Nuclear export of unspliced HIV-1 mRNAs requires the viral export factor Rev, which specifically binds a cis-acting RNA sequence on viral transcripts called Rev response element (RRE). The nuclear export of the Rev-bound viral RNAs is mediated by interactions with the cellular export protein Crm1 and with yet unidentified additional cellular cofactors.[4]

Recently, the DEAD-box ATPase DDX3 was identified as such an essential cofactor for Rev/RRE-mediated HIV-1 RNA export.[5] However, the molecular details of its role in this pathway are yet to be elucidated. DDX3 is a nucleo-cytoplasmic shuttling protein that binds Crm1 and associates to the cytoplasmic side of nuclear pores. Interestingly, stable knockdown of DDX3 in mammalian cells inhibited HIV-1 replication but did not result in significant cytotoxicity, likely because of a redundancy of functions among the numerous DEAD-box proteins in the cellular but not in the viral metabolism.[6] On the basis of such considerations, DDX3 appears to be an attractive antiviral target, but only a paper describing inhibitors of this enzyme has been reported so far, just after the submission of this manuscript.[7]

Although it is not trivial to obtain specific inhibition of a particular DEAD-box helicase by targeting the nucleotide binding-pocket, specific inhibitors have been developed that are directed toward the nucleotide-binding pocket of other ATPases, such as protein kinases, and it may prove a successful way ahead for the DEAD-box helicases.[8]

The X-ray crystallographic structure of human helicase DDX3 in complex with AMP[9] has been used to generate a structure-based pharmacophoric model to be inserted in a computational protocol for the identification of small inhibitors of the ATPase activity of DDX3. In particular, an analysis of the crystal structure allowed us to identify the chemical features responsible for AMP−enzyme interactions, which were in turn codified into a structure-based pharmacophoric model (Figure 1A and Supporting Information). The protein−ligand interactions are described by five hydrogen bond contacts: the C1 amino and the 2′-OH group of AMP are hydrogen bond donors for the backbone carbonyl oxygen of Arg202 and the phenolic oxygen of Tyr200, respectively; the N7 and two phosphate oxygens of AMP are the hydrogen bond acceptors for the side chain NH_2 group of Gln207 and the backbone NH group of Gly229, respectively. In addition, an aromatic ring feature on the phenyl ring of adenine describes $\pi-\pi$ interactions with the aromatic portion of Tyr200. Finally, 10 excluded volumes, localized in region of space where amino acid backbone or side chains lie, identify zones forbidden to the ligands. The resulting model consisted of six pharmacophoric features (two hydrogen bond donors, three hydrogen bond acceptors, and an aromatic ring) that accounted for the interactions between AMP and DDX3, and 10 excluded volumes corresponding to regions of AMP binding site occupied by DDX3 portions and, thus, forbidden to any ligand portion. Next, the pharmacophore was used as the three-dimensional query of a virtual screening approach to filter databases of commercially available compounds to identify chemical scaffolds with putative affinity toward the DDX3 ATP binding site, leading to a ranking list of 70 compounds. Such compounds were docked into the DDX3 ATP binding site, and a consensus scoring based on Gold (version 3.1.1)[10] was applied to improve the enrichment of true positives by the analysis of

* To whom correspondence should be addressed. For G.M.: phone, +39 0382 546354; fax, +39 0382 422286; e-mail, maga@igm.cnr.it. M.B.: phone, +39 0577 234306; fax, +39 0577 234333; e-mail, botta@unisi.it.
† Istituto di Genetica Molecolare.
‡ Università degli Studi di Siena.
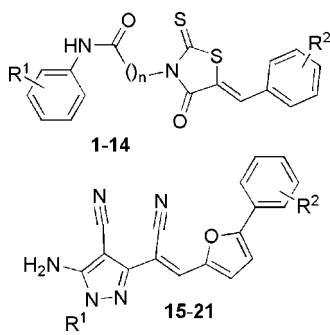§ Katholieke Universiteit Leuven and IRC KULAK.
[a] Abbreviations: DEAD, Aspartic-glutamic-alanine-aspartic; DDX3X, DEAD box polypeptide 3, X-linked; CCR5, Chemokine (C−C motif) receptor 5; CXCR4, Chemokine (C-X-C motif) receptor 4; Gag-Pol: Group Antigens-Polymerase; Env, Envelope; Rev, Regulator of virion protein; RRE, Rev response element; Crm1, chromosome region maintenance 1; HCV NS3, Hepatitis C virus nonstructural region 3.
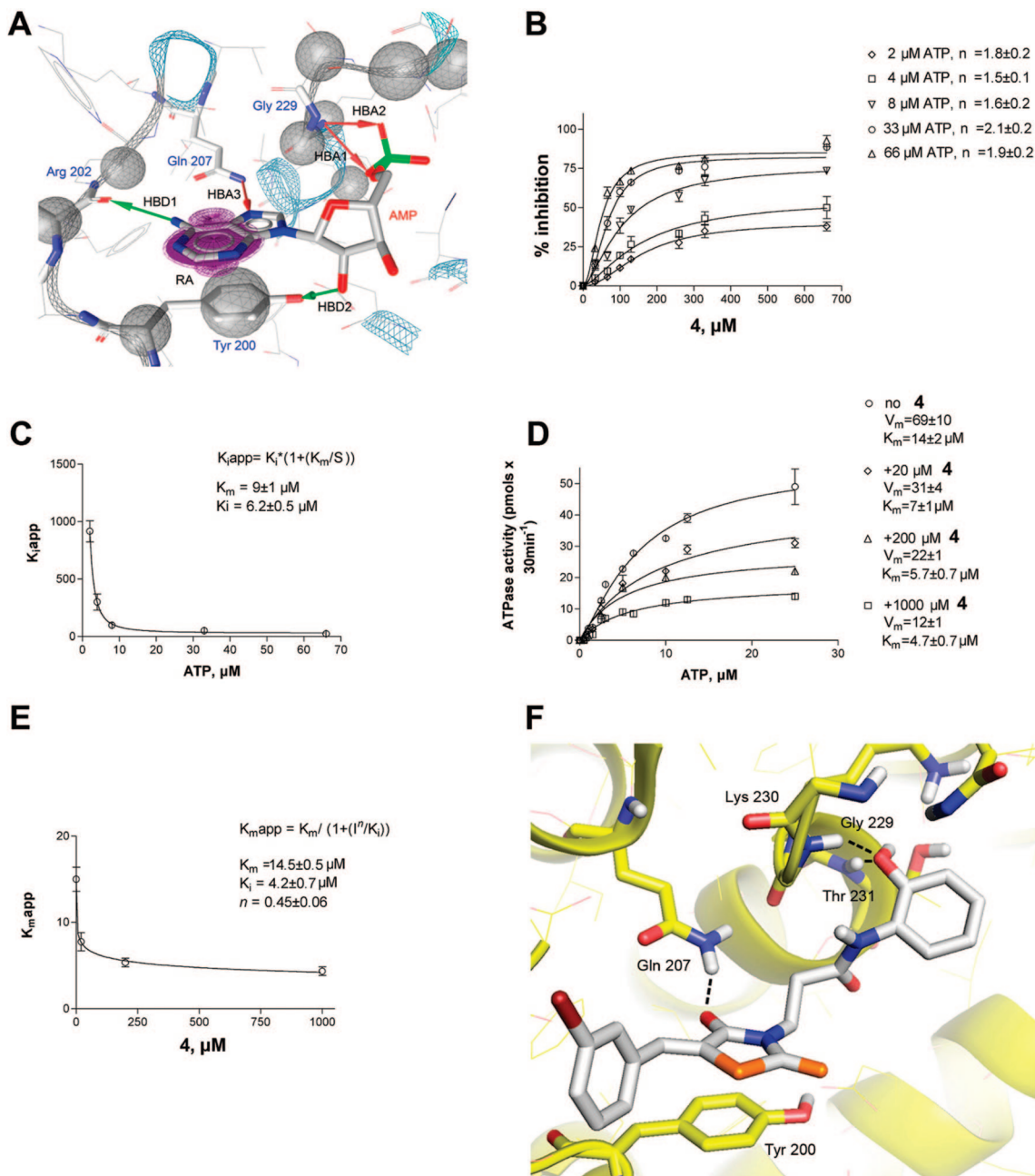
Figure 1. (A) Graphical representation of the structure-based pharmacophoric model built from the crystallographic structure of the DDX3−AMP complex. Ten excluded volumes (gray spheres), three hydrogen bond acceptors (HBA, red arrows), two hydrogen bond donors (HBD, green arrows), and an aromatic ring (RA, violet) constitute the model. (B) Dose-dependent inhibition of the ATPase activity of human DDX3 by **4** at different ATP concentrations. Values are the mean of three independent experiments $\pm$ SD. (C) Variation of the apparent inhibition constant ($K_{i(app)}$) of **4** as a function of the ATP concentration. Values are the mean of three independent experiments $\pm$ SD. (D) Variation of the rate of ATP hydrolysis by DDX3 as a function of the ATP concentration in the absence (circles) or in the presence of increasing amounts of **4** (20 μM, diamonds; 200 μM, triangles; 1000 μM, squares). Values are the mean of three independent experiments $\pm$ SD. (E) Variation of the apparent affinity of DDX3 for the ATP substrate ($K_{m(app)}$) as a function of the **4** inhibitor concentration. Values are the means of three independent experiments $\pm$ SD. (F) The binding mode of compound **4** (atom type notation) within the DDX3 ATP domain found by docking simulations. Black dashed lines represent intermolecular hydrogen bond contacts.

their ability to have profitable interactions with the enzyme and, at the same time, to further prune the previous ranking list. Analysis of docking results, on the basis of the scoring functions of the software or a visual inspection of the docked poses of each compound, led us to choose 10 out of the original 70 entries to be purchased and submitted to in vitro biological tests. Among them, three hit compounds (namely, **1**, **2**, and **15**, Table 1), belonging to two different structural classes and showing IC$_{50}$ values in the range 90−300 μM, were identified and used to generate two small focused libraries of commercially available

compounds. Testing the latter led to the identification of the rhodanine-based derivative **4** with an inhibitory potency of 5.4 μM, significantly improved with respect to parent compounds. The computational model showing the proposed interaction pattern between **4** and the DDX3 ATP binding pocket (Figure 1F) predicted the 2′-OH group of the inhibitor as a hydrogen bond acceptor contacting the backbone NH moiety of Gly229 and Thr231. Moreover, the amide carbonyl oxygen was estimated at hydrogen bond distance from the NH group of Ala232, while the alkyl portion of the propanamide moiety could

**Table 1.** Inhibitory Activity of Compounds **1**−**21** toward the Human Cellular DEAD-Box ATPase DDX3

compd	R^1	R^2	n	activity[a]
1	3-OH	2-OCH$_3$	2	200
2	3-OH	3-Br	2	90
3	4-OH	3-Br	2	300
4	2-OH	3-Br	2	5.4
5	3-OH	3-Br	3	150
6	3-OH	3-Br	1	NA
7	3-OH	H	2	NA
8	2-OH	3-Br	3	500
9	2-OH	H	2	NA
10	2-OH	2-Cl	2	NA
11	2-OH	4-CH$_3$	2	NA
12	2-OH, 3-Cl	4-CH$_3$	2	NA
13	2-COOH	3-Br	2	200
14	2-OCH$_3$	3-Br	2	NA
15	Ph	2-NO$_2$		300
16	(CH$_2$)$_2$OH	2-NO$_2$		NA
17	H	3-NO$_2$		38
18	Ph	3-Cl		500
19	H	2-NO$_2$, 4-CH$_3$		NA
20	H	4-COOH		NA
21	Ph	2-Cl		NA

[a] Expressed as μM concentration. NA: not active at the test dose.

give hydrophobic interactions with Ala232. The rhodanine carbonyl oxygen was suggested to interact with the terminal NH$_2$ of the Gln207 side chain, and π−π interactions between the rhodanine nucleus and the aromatic portion of Tyr200 were hypothesized. Finally, the bromobenzyl terminus was embedded within the hydrophobic cavity delimited by the polymethylenic chains of Glu523 (subunit B), Arg139, and Arg202.

Further biological characterization of **4** showed that increasing amounts of the test compound led to a dose-dependent inhibition of the ATPase activity with a sigmoidal relationship as a function of the inhibitor concentration (Figure 1B). Computer-aided best fitting to the experimental data indicated binding with a cooperativity index (n) significantly higher than 1 (between 1.5 and 2.1, F test against the $n = 1$ hypothesis resulted in $p < 0.002$), suggesting the presence of two nonequivalent binding sites for **4** on DDX3.

Inhibition of ATPase activity by **4** changed as a function of the ATP concentration (Figure 1B and Supporting Information), leading to an apparent inhibition constant ($K_{i(app)}$) decreasing as the ATP concentration increased (Figure 1C). By measurement of the ATPase reaction rate as a function of the ATP concentration in the absence or in the presence of different inhibitor concentrations, the maximal reaction rate (V_m) and the apparent affinity of the enzyme for the ATP substrate ($K_{m(app)}$) were decreased by the inhibitor (Figure 1D,E). This evidence suggested **4** as an uncompetitive inhibitor of human DDX3. The inhibitor dissociation constant (K_i) calculated from the variation of the $K_{i(app)}$ or $K_{m(app)}$, according to the uncompetitive inhibition model, was 6.2 and 4.5 μM, respectively, in good agreement with the IC$_{50}$ of 5.4 μM obtained from inhibition assays in the
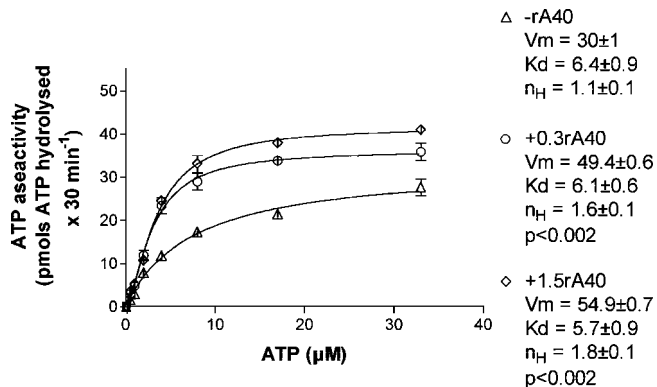
Figure 2. Variation of the DDX3 ATPase activity of human DDX3 as a function of the ATP concentration in the absence (triangles) or in the presence of different amounts (0.3 μM, circles; 1.5 μM, diamonds) of the oligoribonucleotide rA40. Values are the mean of three independent experiments. Error bars represent ±SD. The p values were calculated by an F test against the no cooperativity ($n = 1$) hypothesis.

presence of saturating concentrations of ATP. When tested toward other ATP-dependent enzymes (HCV NS3 ATPase/helicase, T4 polynucleotide kinase, c-Src protein kinase), **4** did not show inhibitory activity when tested up to 100 μM.

On the other hand, **4** inhibited the replication of HIV-1(III$_B$) in MT-4 cells with an EC$_{50}$ of 86.7 μM, without showing cytotoxicity at 125 μM. Also, no toxicity was found in MOLT-4 T-lymphocytic cells up to 200 μM. The antiviral activity of **4** found in MT-4 cells was lower than the inhibitory effect observed toward the ATPase activity of recombinant DDX3. This might be explained by its unfavorable pharmacokinetic properties.

Experimental evidence on the peculiar binding mode of **4** suggested that DDX3 is acting as a multimeric enzyme. This is not surprising, since most cellular and viral DNA and RNA helicases are multimeric enzymes. Binding of a nucleic acid to DDX3 causes cooperative binding of the ATP substrate (Figure 2),[11] likely due to a stabilization of the multimeric form. As a consequence, it is possible that the binding of **4** to the ATP site of a DDX3 monomer stabilizes the multimeric form of the enzyme, causing a conformational change that facilitates binding of a second inhibitor molecule to the next monomer, thus resulting in suppression of the ATPase activity.

A structure−activity relationship of rhodanine analogues **1**−**14** clearly showed that the 2-hydroxy group of **4** is a crucial determinant for affinity (Table 1). In fact, the corresponding methoxy derivative **14** was inactive, while **2** and **3** (bearing the OH substituent at position 3 and 4, respectively) showed affinity 16- and 55-fold lower, respectively. These findings are in agreement with the binding mode proposed for **4**, showing the hydroxy group within the polar region corresponding to the AMP phosphate binding site and engaged in hydrogen bond contacts with the NH group of Gly229 and Thr231.

Moreover, the ethylenic spacer of propanamide was also very important for affinity. In fact, increasing the length of the alkyl spacer from two to three carbon atoms led to **8** with an affinity (500 μM) about 2 orders of magnitude lower than that of **4** (5.4 μM). An analysis of the proposed binding mode revealed that the propanamide spacer of **4** had the optimal length to bring the phenolic and the benzylrhodanine portions of the molecule to the right distance to allow them to make profitable hydrophobic and hydrogen bond contacts with amino acids of the AMP binding site.

Finally, changing of the *m*-Br substitution with various groups in different substitution patterns generated inactive compounds such as **9−12**. This is probably due to the fact that such a substituent and its substitution pattern represent an optimal system to interact with the polymethylene lipophilic portions of the side chains of Glu523, Arg139, and Arg202.

On the other hand, the small number of pyrazole derivatives **15−21** did not allow us to make any structure−activity relationship consideration, although **17** showed an interesting IC$_{50}$ value (38 μM), 7-fold higher with respect to that of the most active compound **4**.

In summary, we report here the discovery of an inhibitor of the human cellular cofactor DDX3 (known to be involved in HIV-1 infection) and demonstrate that targeting this cellular enzyme may stop HIV-1 infection. Considering that **4** is the first small molecule showing inhibitory activity toward HIV-1 replication by targeting the human cellular cofactor DDX3, it represents a hit compound for a completely new class of antiviral agents. A rational drug optimization process has been also planned to synthesize derivatives of **4** with improved activity with respect to the parent compounds.

Supporting Information Available: Further details on computational methods and biological assays; elemental analysis results. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Habeshaw, J. A.; Dalgleish, A. G.; Bountiff, L.; Newell, A. L.; Wilks, D.; Walker, L. C.; Manca, F. AIDS pathogenesis: HIV envelope and its interaction with cell proteins. *Immunol. Today* **1990**, *11*, 418–425.

(2) Fätkenheuer, G.; Pozniak, A. L.; Johnson, M. A.; Plettenberg, A.; Staszewski, S.; Hoepelman, A. I.; Saag, M. S.; Goebel, F. D.; Rockstroh, J. K.; Dezube, B. J.; Jenkins, T. M.; Medhurst, C.; Sullivan, J. F.; Ridgway, C.; Abel, S.; James, I. T.; Youle, M.; van der Ryst, E. Efficacy of short-term monotherapy with maraviroc, a new CCR5 antagonist, in patients infected with HIV-1. *Nat. Med.* **2005**, *11*, 1170–1172.

(3) (a) Lama, J.; Planelles, V. Host factors influencing susceptibility to HIV infection and AIDS progression. *Retrovirology* **2007**, *4*, 52. (b) Cochrane, A. W.; McNally, M. T.; Mouland, A. J. The retrovirus RNA trafficking granule: from birth to maturity. *Retrovirology* **2006**, *3*, 18.

(4) Yedavalli, V. S.; Benkirane, M.; Jeang, K. T. Tat and trans-activation-responsive (TAR) RNA-independent induction of HIV-1 long terminal repeat by human and murine cyclin T1 requires Sp1. *J. Biol. Chem.* **2003**, *278*, 6404–6410.

(5) Yedavalli, V. S.; Neuveut, C.; Chi, Y. H.; Kleiman, L.; Jeang, K. T. Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. *Cell* **2004**, *119*, 381–392.

(6) Ishaq, M.; Hu, J.; Wu, X.; Fu, Q.; Yang, Y.; Liu, Q.; Guo, D. Knockdown of cellular RNA helicase DDX3 by short hairpin RNAs suppresses HIV-1 viral replication without inducing apoptosis. *Mol. Biotechnol.* **2008**, *39*, 231–238.

(7) Yedavalli, V. S.; Zhang, N.; Cai, H.; Zhang, P.; Starost, M. F.; Hosmane, R. S.; Jeang, K. T. Ring expanded nucleoside analogues inhibit RNA helicase and intracellular human immunodeficiency virus type 1 replication. *J. Med. Chem.* **2008**, *51*, 5043–5051.

(8) Bordeleau, M. E.; Mori, A.; Oberer, M.; Lindqvist, L.; Chard, L. S.; Higa, T.; Belsham, G. J.; Wagner, G.; Tanaka, J.; Pelletier, J. Functional characterization of IRESes by an inhibitor of the RNA helicase eIF4A. *Nat. Chem. Biol* **2006**, *2*, 213–220.

(9) Hogbom, M.; Collins, R.; van den Berg, S.; Jenvert, R.-M.; Karlberg, T.; Kotenyova, T.; Flores, A.; Karlsson-Hedestam, G. B.; Holmberg-Schiavone, L. Crystal structure of conserved domains 1 and 2 of the human DEAD-box helicase DDX3X in complex with the mononucleotide AMP. *J. Mol. Biol.* **2007**, *372*, 150–159.

(10) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43−53. Further details are on the Web page http://www.ccdc.cam.ac.uk/products/life_sciences/gold/.

(11) Franca, R.; Belfiore, A.; Spadari, S.; Maga, G. Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity. *Proteins* **2007**, *67*, 1128–1137.