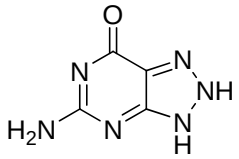
Project: HTS for Identification of splicing modulators at the β-Globin locus

Probe: 8-Azaguanine



SID: 11112293
MLS001074752
Internal IDs: NCGC00016399-01

PubChem Primary Bioassay Identifier (AID): 925

SID	EC50 (μM)	Target	Selectivity*
11112293	2.5	Splicing of Exon 1 of the β-Globin transcript	Unknown

Assigned Assay Grant #: MH080684-01

Screening Center Name: NIH Chemical Genomics Center
Principal Investigator of Screening Center: Christopher Austin

Assay Submitter & Institution: Ryszard Kole, University of North Carolina Dept. Of Pharmacology

Assay or Pathway Target: Splicing of Exon 1 of the β-Globin gene

---

Assay provider information

Specific Aim: Identify small-molecule modulators of β-globin splicing by high throughput screening of chemical libraries.

Significance: β-thalassemia, also known as Cooley's anemia, is a genetic blood disorder caused by a genetic mutation at the β-globin locus. Although over 200 mutations have been revealed to cause β-thalassemia, about 10 mutations are responsible for over 90% of the cases worldwide (1); these include mutations which adversely affect the splicing pattern of the β-globin gene. A mutation in the second intron of β-globin at position 654 (called IVS2-654) leads to the creation of an aberrant 5' splice donor site and activation of a 3' cryptic splice acceptor site (2). This IVS2-654 mutation is frequently found in patients from Thailand and China (3). This form of altered splicing leads to the inclusion of a portion of intron 2 in the mature spliced mRNA, and ultimately results in a prematurely terminated protein. Perturbation of β-globin synthesis affects the ability of red blood cells to carry oxygen, and the severity of the disease depends on the nature of the mutation. Thalassemia

is one of the most common genetic disorders worldwide, and β-thalassemia affects a large number of people in the Middle East, South East Asia, Mediterranean basin, and Africa (4).

Currently, there is no effective drug treatment for β-thalassemia. The ability to target and control pre-mRNA splicing would be an important breakthrough in the search for therapeutics for such diseases. It has been found that splice switching oligonucleotides can manipulate splicing by hybridizing the pre-mRNA at elements critical for splicing factor recognition. This program was designed to use a cell line containing a mutant intron-2 of IVS2-654 inserted into a stably expressed EGFP gene for identifying the small molecule modulators that block the aberrant splicing.

Screening center information

Assay Implementation and Screening

PubChem Bioassay Name: qHTS Assay for Modulators of Hemoglobin Beta Chain Splicing

List of PubChem bioassay identifiers generated for this screening project (AIDs):

AID	Target	Concentration	Bioassay type
1405	Summary	N/A	Summary assay
925	β-Globin Locus ISV2-654 primary screen	46μM – 3 nM	Primary qHTS (membrane potential measurement)
910	Confirmation in ISV2-654	46μM – 3 nM	Followup assay
930	Confirmation in ISV2 705 cell line	46μM – 3 nM	Followup assay
931	Cytotoxicity in ISV2-654	46μM – 3 nM	Cytotoxicity ATP quantitation

Primary Assay Description as defined in PubChem:

Overview:

This assay was designed for screening small molecule modulators of the aberrant RNA splicing in β-thalassemia. The intron-2 of IVS2-654 mutation in human β-globin was inserted into the enhanced green fluorescent protein (EGFP) gene that was expressed stably in a HeLa cell line. The functional EGFP is expressed when the alternative splicing due to the mutation is blocked. One benefit of designing the assay to detect the compound effect on increase in functional EGFP expression is that cytotoxic compounds can be eliminated because they decrease the EGFP signal. The identified small molecule probes were further tested in the RT-PCR assay to confirm the activity on blocking the aberrant splicing in cell lines. In addition, the effect of the small molecules was measured *in vivo* with the disease model of IVS2-654 β-thalassemia (5).

Assay design and protocol:

A HeLa EGFP-654 cell line stably expresses a construct containing the enhanced green fluorescent protein (EGFP) gene that is interrupted by the human IVS2-654 β-globin intron (6) (Fig. 1a). A portion of the intron is inserted into the mRNA in this cell line, which results in an interruption of the translation of EGFP open reading frame and premature termination

of EGFP protein synthesis. The EGFP proteins are not expressed in the cells. On the other hand, when a control antisense oligo blocks the aberrant splicing site in the intron of IVS2-654 gene, the normal mRNA for EGFP protein is formed and the full length EGFP proteins are synthesized, which generate green fluorescence in cells **(Fig. 1b)**. In addition, a second cell line expressing a construct with the intron 2 of ISV2-705 mutation inserted into the EGFP gene was used as a secondary assay to test the hits identified from the HTS with the EGFP-654 cells.

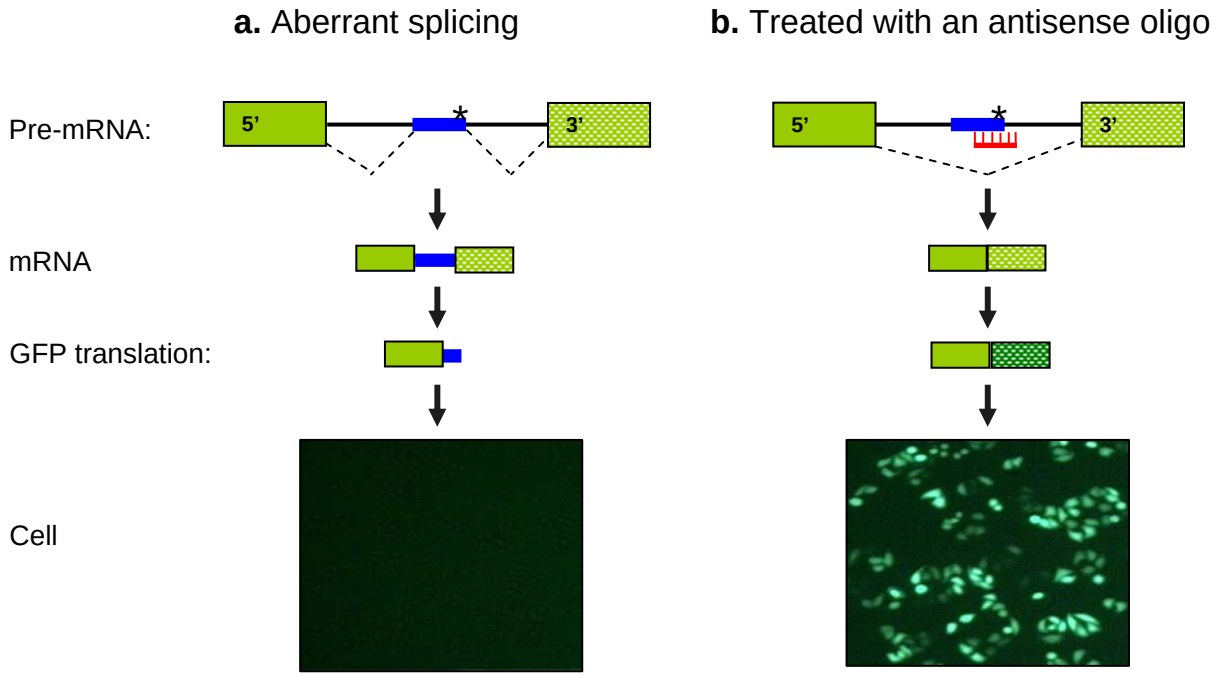


Figure 1. Diagram of assay design for the screening of compounds blocking aberrant splicing due to the intron-2 of IVS2-654 β-globin in β-Thalassemia. (a) Aberrant splicing in the HeLa EGFP-654 cell line due to a mutation in the intron-2 from β-Thalassemia gene that results in an insertion of the intron-2 into mRNA of EGFP. EGFP protein synthesis was interrupted due to a premature stop codon in the intron-2 and thus, no EGFP expression in the cells. (b) In the presence of an antisense oligo that blocked the aberrant splicing site in the intron-2, RNA splicing yielded a full length EGFP mRNA, which was then translated into the normal EGFP proteins in these cells.

The assay method is described in detail in **Table 1**. Briefly, the cells were suspended in the phenol red free Optimem media containing 2% FCS at a density of 100,000 cells/ml. Cells were dispensed at 5 ul/ well into 1536 well black, clear bottom, tissue culture treated plates. After 14 to 16 hours in incubation at 37ºC with 5% CO_2, 23 nl of compound or a control antisense oligo in DMSO solution was added to each well with a pintool station (Kalypsys, San Diego, CA). Cells were further incubated at 37 °C with 5% CO_2 for 40 hours, and the assay plates were then measured in a laser scanning plate cytometer, Acumen Explorer (TTP Labtech, UK), with a laser excitation of 488 nm and emission of 500-530 nm. The fluorescent objects ranging in sizes from 25 to 350 microns in these cells were captured, and a total peak intensity measurement was used for data analysis. The liquid

dispensing, including cells in 1536-well plates, was handled by either by a flying reagent dispensing station (FRD) (Aurora Discovery, San Diego) or a Kalypsys dispensing station (Kalypsys, San Diego, CA).

Table 1. A protocol for the EGFP-ISV2-654 splicing assay

Step	Parameter	Value	Description
1	Reagent	5 μl	500 cells/well
2	Incubation	18 hours	37°C, 5%, CO_2
3	Compound	23 nL	Library compound or control
4	Incubation	40 hours	37°C, 5%, CO_2
5	Detection	Ex=488 and EX=520 nm	Acumen Explorer

Center Summary of the Primary Screen:

The primary screen was performed using this EGFP-ISV2-654 splicing assay against ~65,077 compounds for the compound concentrations ranging from 2.9 nM to 46 mM in a 7-concentration titration.

Data analysis for primary screen results. The primary screen results were analyzed to classify and rank the quality of the concentration-response curves as described previously (Fig. 2) (7). Once an active set of compounds was identified, hierarchical agglomerative clustering with a 0.7 Tanimoto cutoff was performed by using Leadscope (Leadscope Inc., Columbus, OH) fingerprints. For each cluster, maximal common substructures (MCS) were extracted, a manual step of MCS trimming was performed to create a list of scaffolds, and any overlapping scaffolds were abridged to a canonical set. Each scaffold was then represented as a precise definition to indicate descriptors, such as the number of attachment points or the ring size variability. All filters were then relaxed to include the entire negative (class 4) assay data.

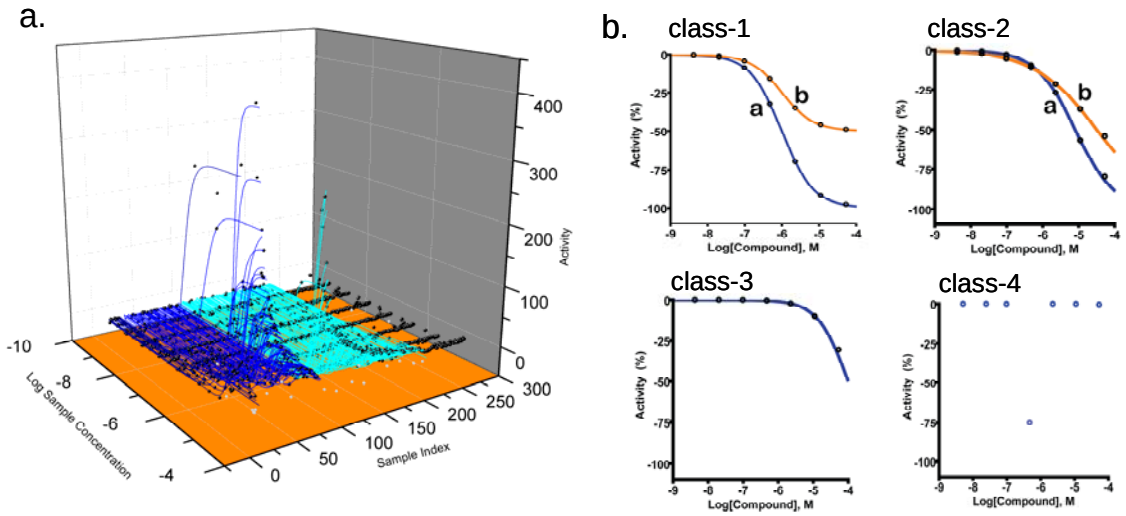


Fig. 2. Example qHTS data and classification scheme for assignment of resulting curve-fit data into classes. **(a)** qHTS curve-fit data from the primary screen binned into curve classifications 1-4 based classification criteria. **(b)** Examples of curves fitting the following classification criteria: Class 1 curves display two asymptotes, an inflection point, and r2 ≥0.9; subclasses 1a vs. 1b are differentiated by full (>80%) vs. partial (≤ 80%) response. Class 2 curves display a single left-hand asymptote and inflection point; subclasses 2a and 2b are differentiated by a max response and r2, >80% and >0.9 or <80% and <0.9, respectively. Class 3 curves have a single left-hand asymptote, no inflection point, and a response >3SD the mean activity of the sample field. Class 4 defines those samples showing no activity across the concentration range.

Hit clustering, SAR analysis and active compound selection: The cluster and SAR analysis of active compounds was carried out based on the primary qHTS data. As all compounds were assayed in at least seven concentrations ranging from 2.94 nM to 46.0 μM, the active compounds were selected by the criteria of potency and quality of concentration-response. 85 compounds (0.1% hit rate) demonstrated a significant concentration-dependent increase in fluorescence, with a curve class of 1.1, 1.2, 1.4, 2.1, or 2.2 (for a detailed definition of curve classification see (ref 7)). Another 182 compounds had activity only at the highest concentration tested (curve class 3). The activities of many of these compounds were not related to a modulation of splicing, but due to the inherent fluorescence of the tested molecules themselves. Even so, the false positive rate for this assay using a laser scanning plate cytometer reader was lower than other fluorescence-based screens that measured the fluorescence intensity in whole-well (8). The fluorescent molecules that preferentially partitioned into live cells were eliminated as the positive compounds by comparing with the data from other cell-based GFP fluorescence assays, including a counter-screen in Chinese hamster ovary cells. A total of 41 compounds were ordered for confirmation and counter-screen assays, as described in the Fig. 3.

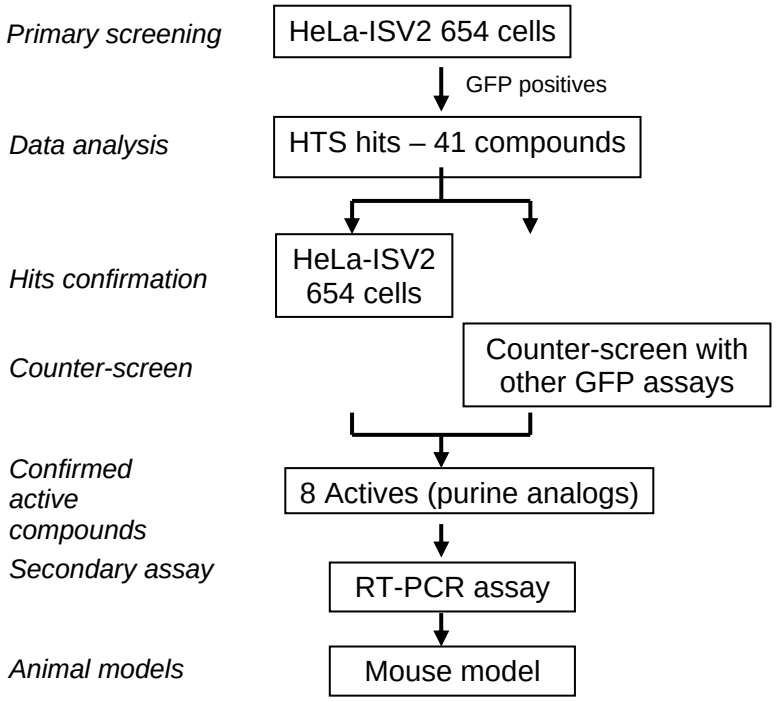


Fig. 3. Flowchart of identification of specific inhibitors of the aberrant splicing caused by a mutation in β-Thalassemia intron by a quantitative high throughput screening.

Confirmation and secondary-screens:

Confirmation and counter-screens. The cherry-picked 41 compounds were tested in the same EGFP-ISV2-654 splicing assay in a titration, as well as in a counter-screen with a CHO cell line expressing EGFP. Eight compounds in a series of purine structure were confirmed, as they were active in the EGFP-ISV2-654 splicing assay but not active in the counter-screen assay.

EGFP-ISV2-705 splicing assay. A stable HeLa cell line harboring the construct with a 705 mutation (9) inserted into the EGFP gene was used as a secondary screen to examine the activity and specificity of the purine analogs. All the active compounds tested exhibited similar potencies and efficacies in both the 654 and 705 cell lines. The selected compounds and their related analogs were then ordered in powder samples for further testing.

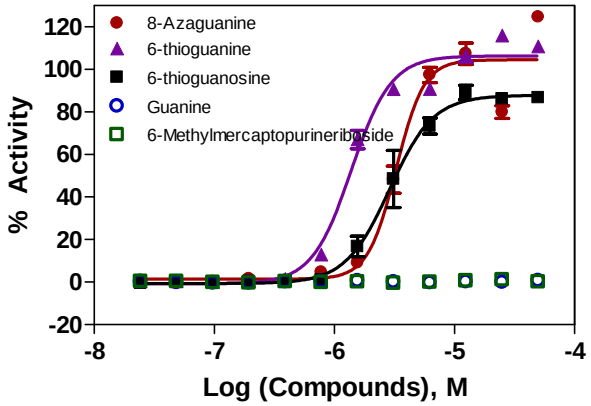


Fig. 4. Concentration responses on correction of aberrant splicing caused by the ISV2-654 mutation by six purine analogs. The EC_{50} values of 8-azaguanine, thioguanine, thioguanosine and 6-Mercaptopurine riboside were 3.3, 1.2, 2.0 and 2.0 μM, respectively. The other two purine analogs, Quanine and 6-Methylmercaptopurine riboside, were not active in this assay.

Cytotoxicity analysis. Cytotoxicity of the compounds was measured in the HeLa EGFP-ISV2-654 cells using an ATP-Lite assay kit (Perkin Elmer Boston, MA). Cells were plated in 1536-well plates and incubated with compounds for the same amount of time as in the primary screen assay. After 40 hour incubation in the presence of compounds, 4 ul of ATP-Lite reagent was added to each well to detect the ATP content in cells. The luminescence signal was measured in a Viewlux plate reader (Perkin Elmer, Boston MA). The reduction of ATP content by compounds indicated the cytotoxicity (Appendix Table 2).

RT-PCR assay. A RT-PCR assay was used to measure the ability of these confirmed compounds to correct the aberrant splicing due to the intron-2 of IVS2-654 mutation. In above experiments, 8-azaguanine was confirmed to correct the aberrant splicing in the EGFP-IVS2-654 cell line in a concentration dependant manner, with an EC_{50} of 3.3 uM (Fig. 4). We found that its activity on correction of aberrant splicing was confirmed by quantifying relative expression of the correctly- and aberrantly-spliced mRNAs with a RT-PCR experiment (Fig. 5a). Furthermore, in another β-globin-IVS2 cell line that contained the full-length mutant β-globin gene, 8-azaguanine was able to correct the aberrant splicing of the thalassemic β-globin transcript as measured by the RT-PCR method (Fig. 5b). Several other purine analogs including 6-thioguanine, 6-thioguanosine and 6-mercaptopurine riboside were also confirmed to be active in the EGFP-IVS2-654 cell line by microscopy and RT-PCR (data not shown).

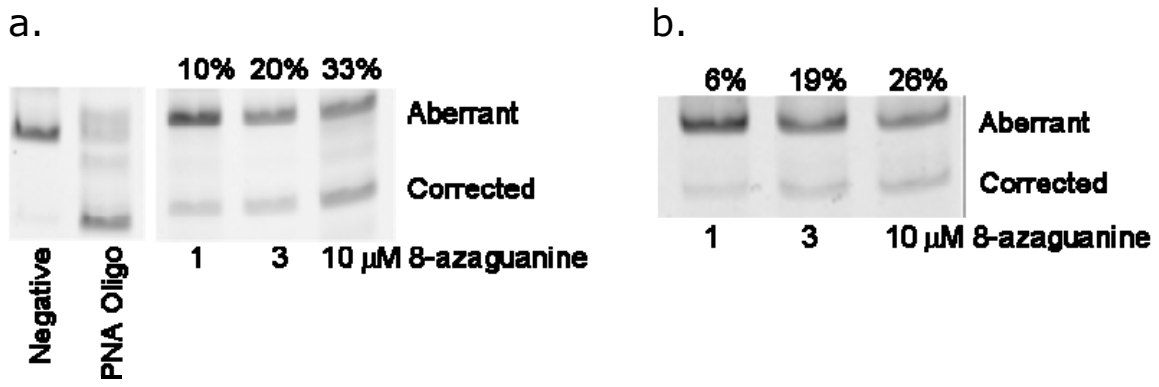


Fig. 5. RT-PCR analysis of splicing modulation by 8-azaguanine. (a) In the EGFP-ISV2-654 cell line, correctly spliced mRNA significantly increased after the treatment with 8-Azaguanine, as determined by RT-PCR experiment. (b) In a β-globin-ISV2-654 cell line, the correctly spliced mRNA significantly increased after treatment with 8-azaguanine.

Probe Characterization

Prior Art for the Splicing Modulators of EGFP-ISV2-654. There have been no reported small molecule modulators that specifically modulate the splicing in the EGFP-ISV2-654 cell line. The oligonucleotides complimentary to the 654 or 705 locus of β-thalassemia have been shown to exhibit effect on the splice site switching specific to the locus (2).

Initial Probes Modulating RNA Splicing in EGFP-ISV2-654 cells. The activity on modulating the aberrant splicing by 8-Azaguanine and several other purine analogs were confirmed in the EGFP-ISV2-654 splicing assay. Compounds were also tested for cytotoxicity in the same cell line with an ATP content assay. Some purine analogs were cytotoxic but had no effect on splicing, indicating these two effects may not be linked. 8-Azaguanine had the low levels of cytotoxicity, while it displayed a significant modulation of the RNA splicing in the EGFP-ISV2-654 cells. 8-Azaguanine was also among the most active compounds in the ISV2-705 secondary assay, and its activity was confirmed by the RT-PCR experiment.

Known probe properties

Center summary of probe properties (solubility, absorbance/fluorescence, reactivity, toxicity, etc.) and recommendations for the scientific use of probe as research tool:

8-Azaguanine
MLS001074752

Molecular Weight	152.11416 [g/mol]
Molecular Formula	$C_4H_4N_6O$
XLogP	-1
H-Bond Donor	3
H-Bond Acceptor	6
Rotatable Bond Count	0
Tautomer Count	23
Exact Mass	152.044659
MonoIsotopic Mass	152.044659
Topological Polar Surface Area	104
Heavy Atom Count	11

Has this compound been provided to the MLSMR?

MLS001074752 is already in the MLSMR library.

Canonical SMILES: C12=NNNC1=NC(=NC2=O)N

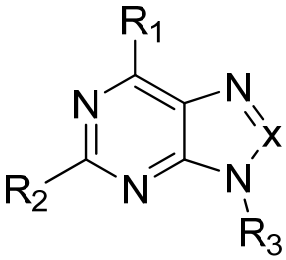
InChI: InChI=1/C4H4N6O/c5-4-6-2-1(3(11)7-4)8-10-9-2/h(H4,5,6,7,8,9,10,11)/f/h9-10H,5H2

Compound preparation: Compound is prepared in DMSO at 10 mM stock concentration.

Probe availability: 8-Azaguanine is available from Sigma-Aldrich (A5284).

Appendices-A

Table 2. SAR for the β-globin splicing probe and its analogues.



Compound	Name	R1	R2	R3	x	EGFP Splicing (uM)	EGFP Response (%)	Cytotoxicity (uM)	RT-PCR
1	8-Azaguanine	OH	NH2	H	N	2.5	103	7.1 (++)	+++
2	8-Azathioguanine	SH	NH2	H	N	-	-	-	
3	6-Thioguanine	SH	NH2	H	CH	1.1	108	1.3 (++)	++
4	6-Thioguanosine	SH	NH2	ribose	CH	1.3	97	2.2 (++)	++
5	2-Aminopurine	H	NH2	H	CH	-	-	-	
6	Purine riboside	H	H	ribose	CH	5	113	10 (+)	++
7	6-Mercaptopurine	SH	H	H	CH	1.6	58	2 (+)	
8	6-Mercaptopurine riboside	SH	H	ribose	CH	1.8	58	2.2 (+)	++
9	6-Methylmercaptopurine	SCH3	H	H	CH	-	-	-	
10	6-Methylmercaptopurine riboside	SCH3	H	ribose	CH	-	-	4 (++)	
11	Azathioprine	(a)	H	H	CH	2	57	2.2 (+)	--
12	2-Amino azathioprine	(b)	NH2	H	CH	2	96	3.2 (+)	++
13	Guanine	OH	NH2	H	CH	-	-	-	
14	Hypoxanthine	OH	H	H	CH	-	-	-	
15	Inosine	OH	H	ribose	CH	-	-	-	
16	Adenine	NH2	H	H	CH	-	-	-	
17	Adenosine	NH2	H	ribose	CH	-	-	-	
18	2-Chloroadenosine	NH2	Cl	ribose	CH	-	-	31.6 (+)	
19	8-Azidoadenosine	NH2	H	ribose	C-N3	-	-	-	--
20	8-Chloroadenosine	NH2	H	ribose	C-Cl	-	-	-	--
21	8-Hydroxyladenosine	NH2	H	ribose	C-OH	-	-	-	--

Bibliography

1. Schwartz E, Benz EJ, Forget BG (1995). Thalassemia Syndromes. In: Hoffman R, Benz Jr. EB, Shattil S, Furie B, Cohen H, Silberstein L, eds. Hematology Basic Principles and Practice, 586-609.

2. Sazani P, Kole R (2003). Therapeutic potential of antisense oligonucleotides as modulators of alternative splicing. *J Clin Invest.* 112(4):481-486.

3. Sierakowska H, Sambade MJ, Agrawal S, Kole R (1996). Repair of thalassemic human β-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc Natl Acad Sci U S A*. 93(23):12840-12844.

4. Rund D, Rachmilewitz E (2005). Medical progress: β-thalassemia. *New England Journal of Medicine* 353(11):1135-1146.

5. Lewis J, Yang B, Kim R, et al (1998). A common human β globin splicing mutation modeled in mice. *Blood*. 91(6):2152-2156.

6. Sazani P, Kang SH, Maier MA, et al (2001). Nuclear antisense effects of neutral, anionic and cationic oligonucleotide analogs. *Nucleic Acids Res* 29(19):3965-3974.

7. Inglese J, Auld DS, Jadhav A, et al (2006). Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc Natl Acad Sci U S A* 103(31):11473-11478.

8. Simeonov S, Jadhav A, Thomas CJ, Wang Y, Huang R, Southall N, Shinn P, Smith J, Austin CP, Inglese J (2008). Fluorescence spectroscopic profiling of compound libraries. *J. Med. Chem.* 51(8): 2363–2371.

9. Schwartz E, Benz, EJ (1995). The thalassemia syndrome. In:Hematology, Basic Principles and Practice New York: Churchill Livingstone, pp 586-610.