Bioorganic & Medicinal Chemistry Letters 20 (2010) 189–192

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

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

Structure-based design, synthesis and biological evaluation of new *N*-carboxyphenylpyrrole derivatives as HIV fusion inhibitors targeting gp41

Yong Wang [a], Hong Lu [b], Qiang Zhu [a,*], Shibo Jiang [b,*], Yun Liao [c,*]

[a] State Key Laboratory of Respiratory Diseases, Guangzhou Institute of Biomedicine and Health, The Chinese Academy of Sciences, Guangzhou 510663, China
[b] Lindsley F. Kimball Research Institute, New York Blood Center, NY 10065, USA
[c] GeometryLifeSci, LLC, 3500 Barrett Dr, # 8M, Kendall Park, NJ 08824, USA

ARTICLE INFO

Article history:
Received 12 October 2009
Revised 29 October 2009
Accepted 30 October 2009
Available online 5 November 2009

Keywords:
HIV
Gp41
Inhibitor

ABSTRACT

A new series of *N*-carboxyphenylpyrrole ligands were designed using GeometryFit based on an X-ray crystal structure of gp41. The synthesized ligands showed significant inhibitory activities against HIV gp41 6-helix bundle formation, HIV-1 mediated cell–cell fusion and HIV-1 replication.

AIDS causes very serious public health problem and economic burden. Globally, an estimated 33 million people are living with HIV, with nearly 7500 new infections each day.[1] So far, 33 anti-HIV drugs (including five fixed-dose combinations) have been approved by FDA with 30 of them belonging to reverse transcriptase inhibitors (RTI) and protease inhibitors (PI).[2] Combination therapies have shown significant synergistic effects.[3] However, there is still no cure or effective vaccine for AIDS, and current anti-HIV drugs are facing increasing prevalence of viral resistance and side effects.[4,5] Therefore, it is in urgent need to develop novel anti-HIV drugs addressing new targets other than RTI and PI.

HIV infects cells through endocytosis and envelope glycoprotein- and dynamin-dependent fusion.[6] Gp41, a transmembrane subunit of HIV-1 envelope glycoprotein, plays a crucial role in HIV fusion and entry.[7] Gp41 exists as a trimer and the ectodomain of each monomer contains an N-terminal peptide (N-peptide) and a C-terminal peptide (C-peptide). HIV fusion involves the insertion of the gp41 N-peptides into host cell membrane and the subsequent binding of the C-peptides (anchored to viral membrane) to N-peptides, which forms a 6-helical bundle bringing the viral membrane and host cell membrane to proximity for fusion. Gp41 is a proven drug target—the peptides derived from the gp41 C-peptides were found as potent HIV fusion inhibitors able to block the N-peptides from binding to the C-peptides.[8–13] One of the peptides, T-20 (Fuzeon/Enfuvirtide, a 36-amino acid synthetic peptide)[14–16] was approved by FDA in 2003 for treating patients with multi-drug resistant HIV. However, T-20 is not orally available and has high production cost.[17] Therefore, developing orally available nonpeptide small-molecule fusion inhibitors targeting gp41 is highly desirable.

So far, no small-molecule anti-HIV drug targeting gp41 has been successfully developed, and gp41 has long been questioned for its druggability for (1) It is difficult for small molecules to block the very strong protein–protein interactions between gp41 C-peptides (red, Fig. 1) and N-peptides (blue, Fig. 1); (2) gp41 N-peptide bundle is highly hydrophobic and lack of deep pockets to allow strong binding with small molecules, which makes it an even more difficult drug target.

Despite the formidable challenges, a few small-molecule gp41 inhibitors with modest activities at micromolar concentrations were discovered during the past 10 years mostly via high-throughput screening and diverse compound library synthesis.[18–25] **A12**

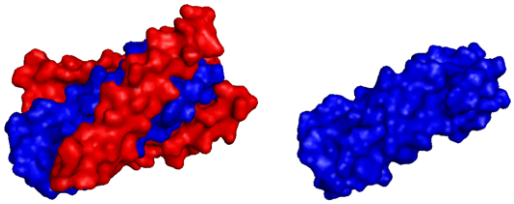


Figure 1. gp41 C- and N-peptide bundle (left); gp41 N-peptide bundle (right). Protein source: 1AIK (PDB code).[8]

* Corresponding authors.
E-mail addresses: zhu_qiang@gibh.ac.cn (Q. Zhu), SJiang@NYBloodCenter.org (S. Jiang), leonard_liao@geometrylifesci.com (Y. Liao).

0960-894X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2009.10.139

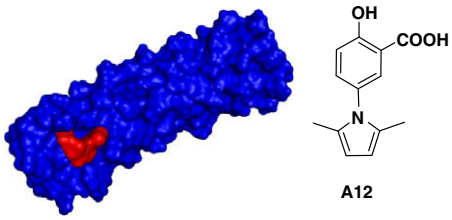


Figure 2. A12 (red) docked into gp41 N-peptide bundle (blue). Protein source: 2R5D (PDB code).[28]

(Fig. 2) is one of the most promising compounds discovered by Xie and Jiang's labs recently.[26] Notably, **A12** is not only able to inhibit gp41 6-helical bundle formation (EC_{50} 37.36 μM) but also druglike.

Until Aug 2009, over 130 X-ray/NMR 3D structures related to gp41 have been released from the Protein Data Bank (PDB),[8,27–29] which provide invaluable information to understand gp41's function at the atomic level as well as lay a solid foundation for structure-based design of small-molecule gp41 inhibitors.

However, structure-based de novo design of small-molecule gp41 inhibitors remains extremely challenging largely due to the lack of effective de novo drug design methodologies.[30] Hamilton's pioneer work[31] in de novo design of gp41 surface antagonists led to a novel small-molecule gp41 inhibitor (EC_{50} ~15 μM), however, its further optimization is difficult partly because its design was based on general α-helix mimicry of gp41 rather than a specific binding pocket on gp41. Most recently, Liu et al. successfully designed a novel small-molecule gp41 inhibitor (IC_{50} 31 μM, inhibition of gp41 helical bundle formation) using their fragment-based de novo ligand design technology, unfortunately, this compound failed to show activity in a HIV mediated cell–cell fusion inhibitory assay in the real life.[32]

We believe that **A12** (Fig. 1) could serve as a valuable starting point for further drug design targeting gp41 by modifying this structure using GeometryFit, a proprietary knowledge-oriented, computer-aided and structure-based de novo drug design methodology developed by GeometryLifeSci.[33]

Based on our docking study (*gp41 N-peptide source*:[28] *PDB ID*: 2r5d) using AutoDock 4.0,[34] we identified a binding mode of **A12** (Fig. 3, ***A12****: pink sticks*) quite different from that proposed previously by Xie and Jiang.[26] In our proposed binding mode, the calculated binding free energy of ΔG is −6.5 kcal/mol and the three key contact motifs identified within the gp41 binding pocket are (1) the flexible and positively-charged Lys38 (*cyan sticks*), which could form a strong salt bridge with the negatively charged carboxylic group of **A12**; (2) the hydrophilic Gln41 (*green*), which could form two H-bonds with the same carboxylic group of **A12**; (3) the highly

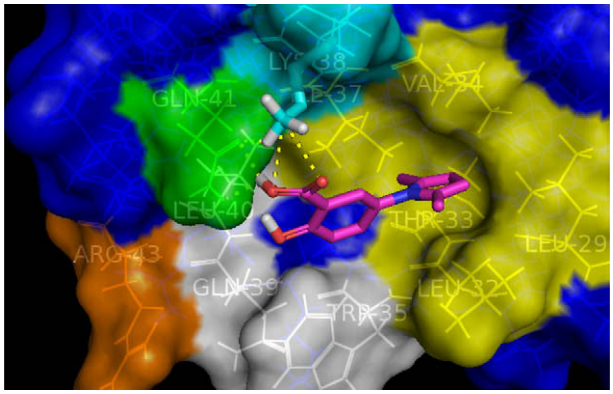
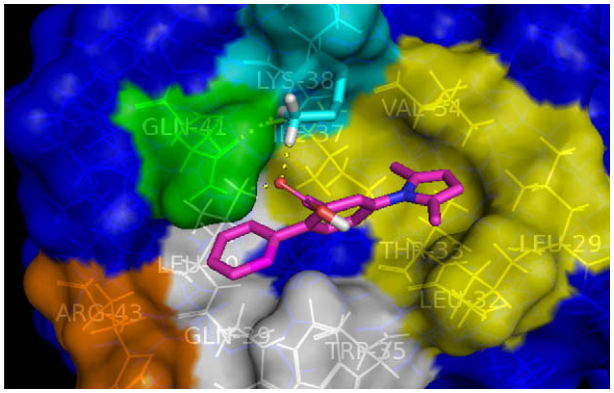


Figure 3. Calculated most favorable binding mode of **A12**.



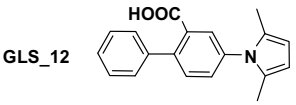
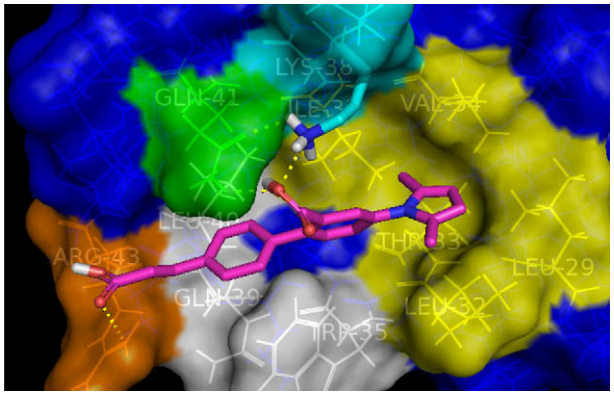


Figure 4. Calculated most favorable binding mode of **GLS_12**.

hydrophobic pocket (*yellow*) composed of Leu29, Leu32, Thr33, Val34 and Ile37, which happens to nicely complement the shape of the hydrophobic pyrrole ring of **A12** to form very favorable hydrophobic interactions. These three key contact motifs may account for the binding of **A12** to gp41, leading to the inhibition of gp41 6-helix bundle formation.

In addition to the three key contact motifs, we then explored additional potential contact motifs within the binding pocket in order to find new ligands with improved binding affinities. The most favorable binding mode of **A12** suggests that its phenolic group do not contribute significantly to its overall binding affinity but do have plenty of room to extend into additional contact motifs within the binding pocket, such as Arg43 (*orange,* Fig 3) and the binding pocket composed of Trp35, Gln39, Leu40 (*white,* Fig. 3).

Using GeometryFit, we designed a new ligand **GLS-12** (Fig. 4) based on the **A12** structure, in which the phenolic group of **A12** was replaced by a phenyl group to not only fit into the shape of the pocket composed of Trp35, Gln39, Leu40 (*white,* Fig. 4) and Gln41 (*green,* Fig. 4) but also generate favorable hydrophobic interactions with Trp35, Gln39, Leu40 (*white,* Fig. 4). The subsequent docking study confirmed that, in the most favorable binding mode of **GLS-12** (Fig. 4), the new ligand well maintains the interactions



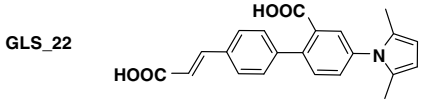
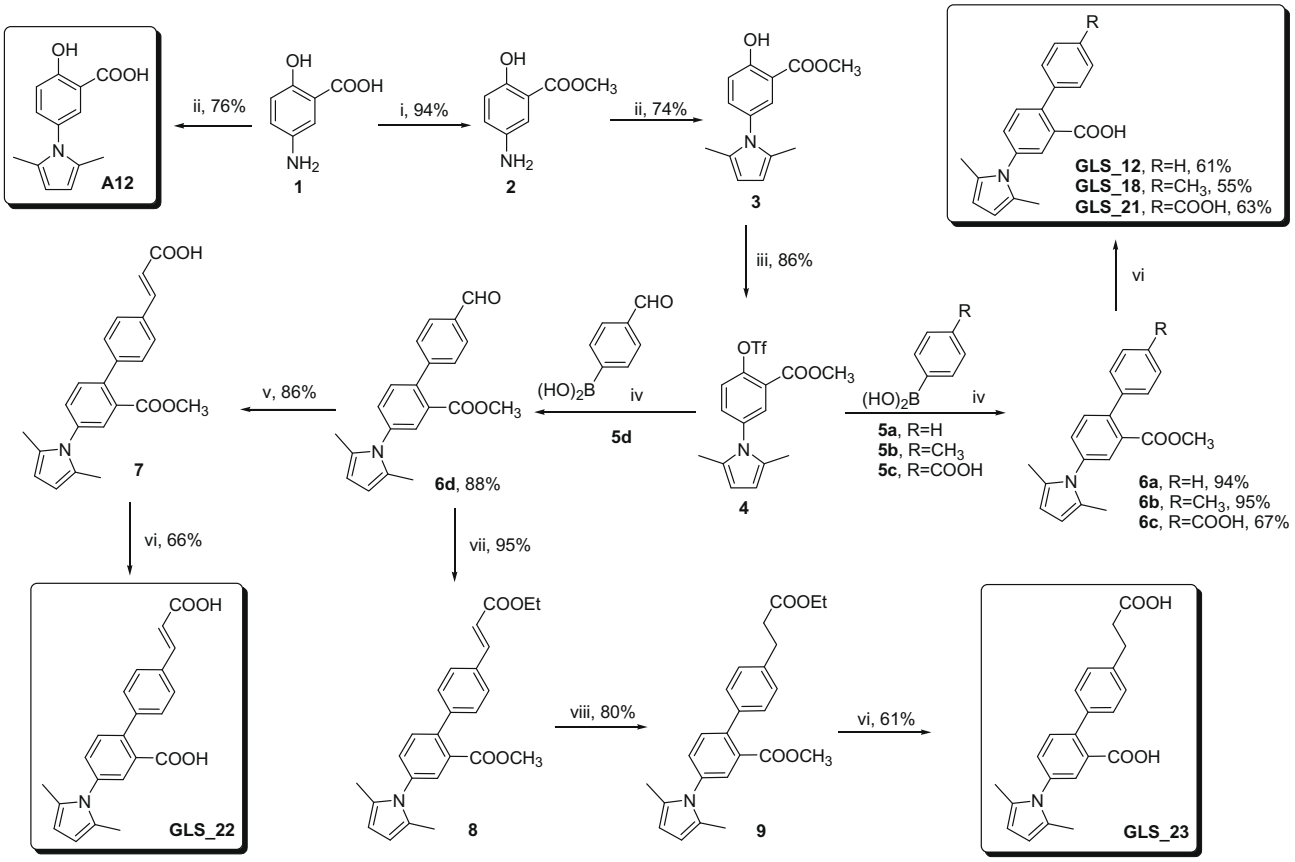


Figure 5. Calculated most favorable binding mode of **GLS_22**.



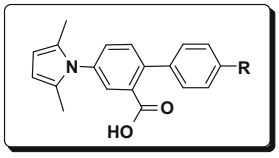
Scheme 1. Reagents and conditions: (i) H_2SO_4, MeOH, reflux; (ii) 2,5-hexanedione, *p*-TsOH, Toluene, reflux; (iii) Tf_2O, Et_3N, DCM, 0 °C; (iv) $PdCl_2(PPh_3)_2$, K_2CO_3, THF/H_2O (2:1), reflux; (v) malonic acid, piperidine, pyridine, reflux; (vi) (1) NaOH, EtOH, (2) HCl; (vii) ethyl (triphenylphosphoranyliden)acetate, DCM; (viii) Pd/C, H_2, THF/MeOH (4:1).

with the previous three key contact motifs while docking nicely into the new binding pocket with its phenyl group. The calculated binding free energy of ΔG was increased to −7.0 kcal/mol.

To further increase the binding affinity, we decided to recruit the positively charged Arg43 (*orange*, Fig. 4) as an additional contact motif within the binding pocket. Using GeometryFit, we designed another new ligand **GLS-22** (Fig. 5) based on the structure of **GLS-12**, in which a negatively charged carboxylic acid handle was added at the *para* position of the phenyl ring in order to form a strong electrostatic interaction with Arg43. The subsequent docking study confirmed that, in the most favorable binding mode of **GLS-22** (Fig. 5), the new ligand's carboxylic group could form a salt bridge with Arg43 and at the same time well maintain the interactions with the previous four key contact motifs. As a result, the calculated binding free energy of ΔG was increased to −8.9 kcal/mol.

In this preliminary study, total of five new ligands were designed, synthesized and screened for their inhibitory activities against (1) HIV gp41 six-helix bundle (6-HB) formation, (2) HIV-1-mediated cell-cell fusion (CF) and (3) HIV-1 replication (p24 and CPE), as previously described,[21,26] using **A12** as a control. The synthesis is illustrated in Scheme 1 and the bio-assay results are summarized in Table 1.

Table 1
Anti-HIV bioassay results versus calculated binding ΔG



GLS_12 R = H
GLS_18 R = Me
GLS_21 R = COOH
GLS_22 R = CH=CHCOOH
GLS_23 R = CH_2CH_2COOH

Ligand[35]	ΔG (calcd) (kcal/mol)	IC_{50} (μM)				CC_{50}[a] (μM)	SI[b]
		p24	CPE	CF	6-HB		CC_{50}/IC_{50} (p24)
A12	−6.5	28.19 ± 3.79	55.45 ± 16.59	43.24 ± 2.28	29.39 ± 5.47	333.84 ± 22.88	11.84
GLS-12	−7.0	17.76 ± 2.74	37.92 ± 7.55	59.17 ± 2.00	28.96 ± 2.27	289.80 ± 13.18	16.31
GLS-18	−6.9	18.19 ± 2.04	79.54 ± 8.16	51.61 ± 3.38	21.26 ± 2.88	280.04 ± 38.11	15.39
GLS-21	−8.7	25.63 ± 1.92	33.93 ± 6.66	26.01 ± 1.31	20.50 ± 1.11	355.23 ± 24.47	13.85
GLS-22	−8.9	4.91 ± 0.69	7.71 ± 1.64	3.60 ± 0.27	20.73 ± 2.11	255.28 ± 3.73	51.99
GLS-23	−8.6	8.30 ± 1.34	13.08 ± 3.19	8.30 ± 0.11	21.75 ± 0.75	227.27 ± 22.86	27.38

[a] CC_{50}: 50% cytotoxicity concentration.
[b] SI: selectivity index.

All 5 designed new ligands showed better activities than **A12** in all assays except **GLS_18** in the CPE assay and **GLS_12** & **18** in the CF assay. The assay results are in accordance with our predictions using GeometryFit in most cases. The designed new ligands (**GLS_21,22,23**) targeting two more contact motifs within the binding pocket than those of **A12** all showed better activities across all four assays, whereas the ligands (**GLS_12,18**) targeting only one more hydrophobic contact motif than those of **A12** mostly showed comparable activities, which indicates that the second additional contact motif, Arg43, may contribute more significantly to the increased binding affinity. The best ligand, **GLS_22**, showed ~6–7-folds better inhibitory activity than **A12** in the HIV replication assays, ~12-folds in the HIV cell–cell fusion assay and ~1.4-folds in the gp41 six-helix bundle formation assay. All five new ligands had low cytotoxicity (CC50: 227–355 μM), and the most potent ligand, **GLS_22**, had the highest Selectivity Index (SI) (CC50/IC50 (p24): 51.99).

Summary: In this preliminary study, we identified a new binding mode of **A12** involving three key contact motifs within the binding pocket, based on which we explored two additional potential contact motifs within the binding pocket and designed five new derivatives of **A12** using GeometryFit. All five new ligands showed improved anti-HIV activities in almost all assays, which suggests that our design model and methodology are reliable, paving the way for our de novo design of novel small-molecule HIV inhibitors targeting gp41 in the future.

Acknowledgement

This work was supported by GeometryLifeSci, LLC, an NIH grant (AI46221) and grant funding from National Science Foundation of China (20942001) and Start-up Foundation for New Investigators from Guangzhou Institute of Biomedicine and Health (GIBH).

References and notes

1. World Health Statistics; WHO, 2009, p 11.
2. http://www.hivandhepatitis.com/hiv_and_aids/hiv_treat.html.
3. De Clercq, E. *J. Clin. Virol.* **2004**, *30*, 115.
4. Richman, D. D.; Morton, S. C.; Wrin, T.; Hellmann, N.; Berry, S.; Shapiro, M. F.; Bozzette, S. A. *AIDS* **2004**, *18*, 1393.
5. (a) Carr, A.; Cooper, D. A. *Lancet* **2000**, *356*, 1423; b Johnson & Johnson strengthens warning on HIV drug, Wed Aug 26, 2009 5:32 pm EDT. http://www.reuters.com/article/marketsNews/idINN26283555200908 26?rpc=44.
6. (a) Miyauchi, K.; Kim, Y.; Latinovic, O.; Morozov, V.; Melikyan, G. B. *Cell* **2009**, *137*(3), 433. and references cited therein; (b) Fackler, O. T.; Peterlin, B. M. *Curr. Biol.* **2000**, *10*, 1005.
7. Moore, J. P.; Jameson, B. A.; Weiss, R. A.; Sattentau, Q. J. In *Viral Fusion Mechanisms*; Bentz, J., Ed.; CRC Press: Boca Raton, FL, 1993; pp 233–289.
8. Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. *Cell* **1997**, *89*, 263.
9. Chan, D. C.; Kim, P. S. *Cell* **1998**, *93*, 681.
10. Jiang, S.; Zhao, Q.; Debnath, A. K. *Curr. Pharm. Des.* **2002**, *8*, 563.
11. Jiang, S.; Lin, K.; Strick, N.; Neurath, A. R. *Nature* **1993**, *365*, 113.
12. Wild, C. T.; Shugars, D. C.; Greenwell, T. K.; McDanal, C. B.; Matthews, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 9770.
13. Lu, M.; Blacklow, S. C.; Kim, P. S. *Nat. Struct. Biol.* **1995**, *2*, 1075.
14. Kilby, J. M.; Eron, J. J. *N. Eng. J. Med.* **2003**, *348*, 2228.
15. Lalezari, J. P.; Henry, K.; O'Hearn, M.; Montaner, J. S.; Piliero, P. J.; Trottier, B.; Walmsley, S.; Cohen, C.; Kuritzkes, D. R.; Eron, J. J., Jr.; Chung, J.; DeMasi, R.; Donatacci, L.; Drobnes, C.; Delehanty, J.; Salgo, M. *N. Eng. J. Med.* **2003**, *348*, 2175.
16. Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. C. *Nature* **1997**, *387*, 426.
17. Liu, S.; Wu, S.; Jiang, S. *Curr. Pharm. Des.* **2007**, *13*, 143.
18. Kazmierski, W. M.; Kenakin, T. P.; Gudmundsson, K. S. *Chem. Biol. Drug Des.* **2006**, *67*, 13.
19. Debnath, A. K.; Radigan, L.; Jiang, S. *J. Med. Chem.* **1999**, *42*, 3203.
20. Zhao, Q.; Ernst, J. T.; Hamilton, A. D.; Debnath, A. K.; Jiang, S. *AIDS Res. Hum. Retroviruses* **2002**, *18*, 989.
21. Jiang, S.; Lu, H.; Liu, S.; Zhao, Q.; He, Y.; Debnath, A. K. *Antimicrob. Agents Chemother.* **2004**, *48*, 4349.
22. Jin, B. S.; Lee, W. K.; Ahn, K.; Lee, M. K.; Yu, Y. G. *J. Biomol. Screen.* **2005**, *10*, 13.
23. Xu, Y.; Lu, H.; Kennedy, J. P.; Yan, X.; McAllister, L. A.; Yamamoto, N.; Moss, J. A.; Boldt, G. E.; Jiang, S.; Janda, K. D. *J. Comb. Chem.* **2006**, *8*, 531.
24. Liu, S.; Lu, H.; Zhao, Q.; He, Y.; Niu, J.; Debnath, A. K.; Wu, S.; Jiang, T. S. *Biochim. Biophys. Acta* **2005**, *1723*, 270.
25. Jiang, S.; Debnath, A. K.; Lu, H. U.S. Patent 0287319 A1, 2006.
26. Liu, K.; Lu, H.; Hou, L.; Qi, Z.; Teixeira, C.; Barbault, F.; Fan, B.-T.; Liu, S.; Jiang, S.; Xie, L. *J. Med. Chem.* **2008**, *51*(24), 7843.
27. (a) Tan, K.; Liu, J.; Wang, J.; Shen, S.; Lu, M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*(23), 12303; (b) Shu, W.; Ji, H.; Lu, M. *J. Biol. Chem.* **2000**, *275*(3), 1839; (c) Shu, W.; Liu, J.; Ji, H.; Radigen, L.; Jiang, S.; Lu, M. *Biochemistry* **2000**, *39*(7), 1634; (d) Wang, S.; York, J.; Shu, W.; Stoller, M. O.; Nunberg, J. H.; Lu, M. *Biochemistry* **2002**, *41*(23), 7283.
28. Welch, B. D.; VanDemark, A. P.; Heroux, A.; Hill, C. P.; Kay, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16828.
29. http://www.rcsb.org/pdb.
30. Mauser, H.; Guba, W. *Curr. Opin. Drug Disc. Dev.* **2008**, *11*(3), 365.
31. Ernst, J. T.; Kutzki, O.; Debnath, A. K.; Jiang, S.; Lu, H.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2002**, *41*(2), 278.
32. Liu, B.; Joseph, R. W.; Dorsey, B. D.; Schiksnis, R. A.; Katrina, N.; Bukhtiyarova, M.; Springman, E. B. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5693.
33. http://geometrylifesci.com.
34. Morris, G.; Goodsell, D.; Halliday, R.; Huey, R.; Hart, W.; Belew, R.; Olson, A. *J. Comput. Chem.* **1998**, *19*, 1639.
35. *Spectral data of the ligands synthesized*: **A12**: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 10.46 (br, 1H), 7.80 (s 1H), 7.38 (d, J = 8.4 Hz, 1H), 7.11 (d, J = 8.8 Hz, 1H), 5.91 (s, 2H), 2.04 (s, 6H). MS: $[M+1]^{+\cdot}$ 232 (ES+APCI). **GLS_12**: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.80 (s 1H), 7.46–7.38 (m, 7H), 5.93 (s, 2H), 2.17(s, 6H). MS: $[M+1]^{+\cdot}$ 292 (ES+APCI). **GLS_18**: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.78 (s 1H), 7.46 (d, J = 8.0 Hz, 1H), 7.39 (d, J = 8.0 Hz, 1H), 7.30–7.22 (m, 4H), 5.93 (s, 2H), 2.41 (s, 3H), 2.09(s, 6H). MS: $[M+1]^{+\cdot}$ 306 (ES+APCI). **GLS_21**: ^{1}H NMR (DMSO-d_6, 400 MHz) δ 7.9 (d, J = 8.0 Hz, 2H), 7.60–7.51 (m, 5H), 5.85 (s, 2H), 2.03 (s, 6H). MS: $[M+1]^{+\cdot}$ 336 (ES+APCI). **GLS_22**: ^{1}H NMR (DMSO-d_6, 400 MHz) δ 7.75 (d, J = 8.0 Hz, 2H), 7.64 (d, J = 16.0 Hz, 1H), 7.56–7.51 (m, 3H), 7.45 (d, J = 8.4 Hz, 2H), 6.58 (d, J = 16.0 Hz, 1H), 5.84 (s, 2H), 2.02 (s, 6H). MS: $[M+1]^{+\cdot}$ 362 (ES+APCI). **GLS_23**: ^{1}H NMR (DMSO-d_6, 400 MHz) δ 7.52–7.47 (m, 3H), 7.40–7.28 (m, 4H), 5.84 (s, 2H), 2.88 (t, 2H), 2.59 (t, 2H), 2.09(s, 6H). MS: $[M+1]^{+\cdot}$ 364 (ES+APCI).