Chemical biology at the crossroads of molecular structure and mechanism

Jennifer A Doudna

Chemical insight into biological function is the holy grail of structural biology. Small molecules are central players as building blocks, effectors and probes of macromolecular structure and function.

The recent explosion of interest in chemical biology reflects fascination with research at the interface of chemistry and biology, where chemical insights are brought to bear on exciting biological and biomedical problems. In truth, however, small molecules have long had an important role in deciphering the answers to fundamental biological questions. For example, the use of small-molecule modifiers of DNA, and later dideoxy nucleoside triphosphates, provided the first insights into the sequence of DNA and led to development of the fluorescent nucleotide analogs that enabled whole-genome sequencing. Early small-molecule indicators for intracellular calcium led to the discovery of the calcium concentration changes that signal everything from egg fertilization to memory formation.

More recently, small molecules have emerged in starring roles as tools for determining macromolecular structures and probing molecular interactions and dynamics. In addition, as discussed in an earlier Commentary in this series[1], small molecules participate directly in many biological processes, functioning in diverse roles as chemical triggers, inhibitors, stimulators and switches. Thus, understanding how biological macromolecules work in chemical detail often requires both clarifying the native functions of small-molecule ligands and designing new ones as probes. Although frequently functioning as covalent and noncovalent effectors of macromolecular structure and function, small molecules are also the monomeric starting materials for biosynthesis of macromolecules. The ability to incorporate chemically modified versions of these building blocks into biopolymers including DNA, RNA, oligosaccharides and polypeptides by manipulating synthetic and biosynthetic pathways has revolutionized our ability to produce and study biopolymers *in vitro* and *in vivo*.

Small molecules thus provide a common link between the fields of chemical, structural and cell biology. The major expansions in the fields of RNA biology and intracellular signaling pathways in recent years, in

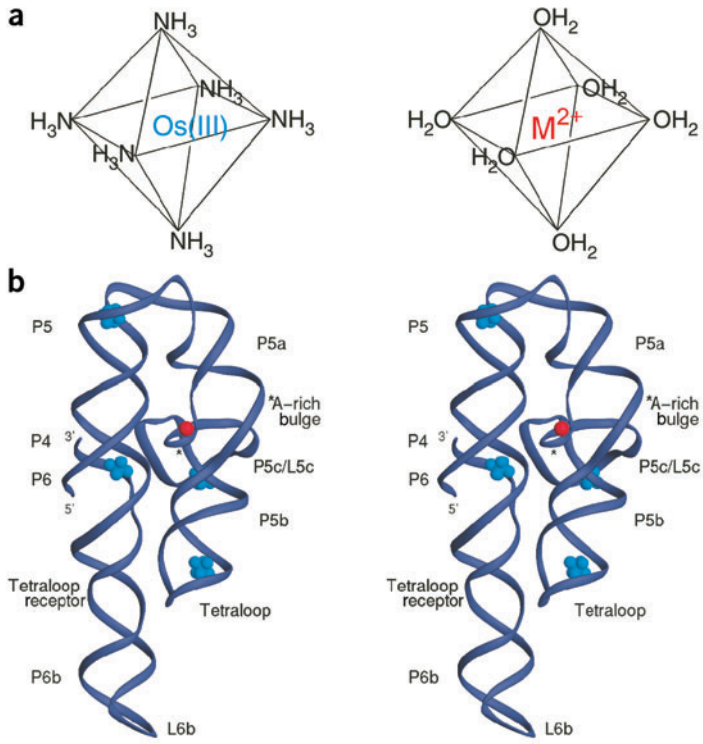


Figure 1 Osmium hexammine is a structural mimic of hexahydrated magnesium ion. (**a**) Osmium(III) hexammine and divalent ions such as magnesium and manganese share octahedral coordination geometry. (**b**) Stereo image of the *Tetrahymena* P4–P6 domain, with osmium hexammine binding sites shown in light blue and an inner-sphere magnesium binding site shown in red[3]. Reprinted from ref. 3 with permission from Elsevier.

Jennifer A. Doudna is in the Howard Hughes Medical Institute, Department of Molecular and Cell Biology and the Department of Chemistry, University of California, Berkeley, Berkeley, California 94705, USA.
e-mail: doudna@berkeley.edu

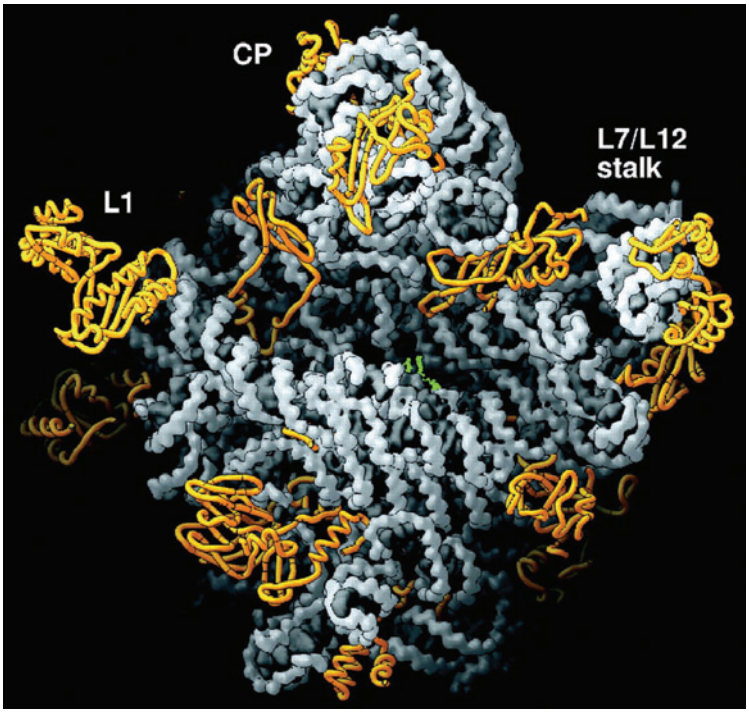


Figure 2 Crystal structure of the *Haloarcula marismortui* 50S ribosomal subunit[5]. The all-RNA active site of the ribosome is a key binding site for antimicrobial small molecules. Reprinted with permission from ref. 5. © 2000 American Association for the Advancement of Science (AAAS).

particular, have been fueled by the potent mix of structural biology combined with genetic manipulation, insights from small-molecule effectors, and advances in chemical and enzymatic synthesis.

Hexammines: an RNA crystallographer's best friend

The journal *Inorganic Chemistry* might not be the first place one would look for insights into the structure and function of large RNA molecules. But it was an article published there in 1989 that provided the key to solving the first crystal structure of a large RNA, the P4–P6 domain of the *Tetrahymena thermophila* group I intron. Taube and colleagues described a method for preparing osmium hexammine[2], a nearly perfect geometric mimic of magnesium hexahydrate, which turned out to bind to particular sites in the major groove of RNA double helices[3] (**Fig. 1**). Because osmium is electron-rich relative to the atoms found in biological molecules, it served as an ideal metal with which to derivatize the P4–P6 RNA crystals and determine the structure factor phases essential for calculating an electron density map[4].

Binding of osmium hexammine to a few specific sites might have been a lucky but unusual property of the P4–P6 RNA. However, hexammines of osmium, iridium and cobalt, which tend to bind at the major-groove side of tandem G-U base pairs, have subsequently been used to solve numerous RNA and RNA-protein crystal structures, including the ribosomal subunits[5,6], the intact ribosome[7], a guanine-responsive riboswitch[8] and ribonuclease P (refs. 9,10). In each case, hexammine binding sites within the RNA occur with a frequency of one per 30–50 nucleotides, providing plenty of detectable signal in crystallographic experiments without perturbing the native conformation of the molecules in the crystal.

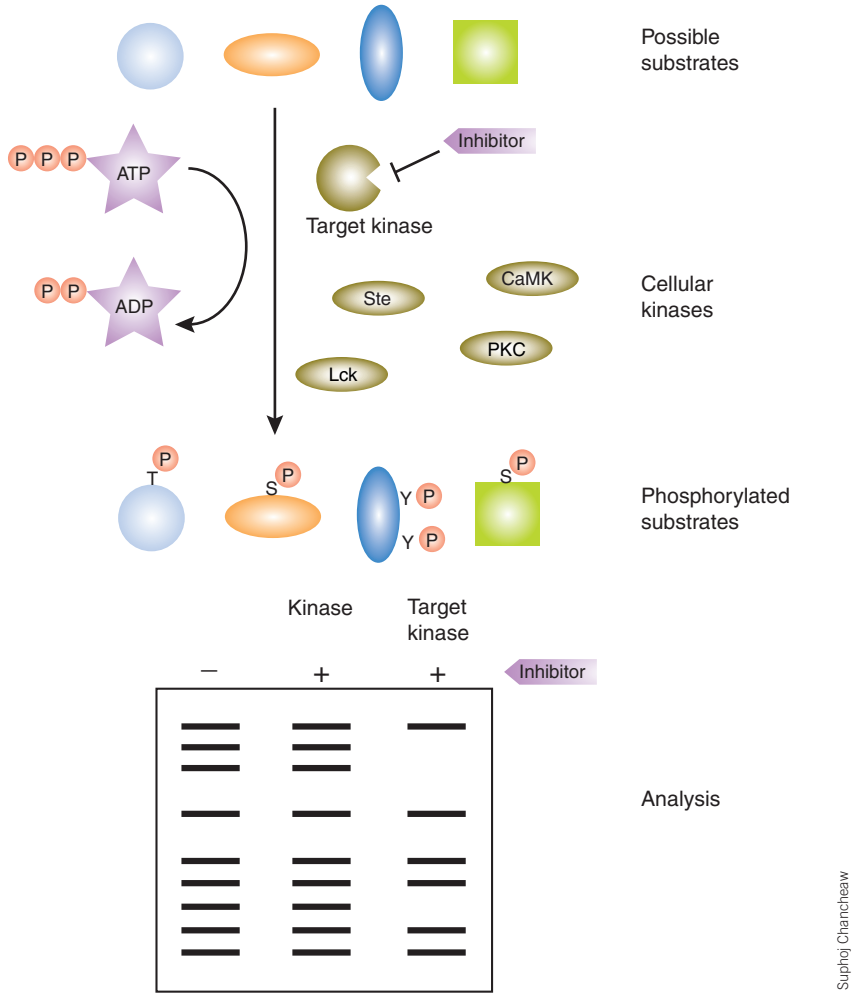
Although hexammines are useful derivatives and probes for magnesium hexahydrate binding sites in RNA, similar chemical mimics are lacking for the classes of partially hydrated magnesium ion binding sites that are often critical for RNA folding and function. Although pentaammines offer some utility for detecting and perturbing magnesium ions that have a single inner-sphere coordination site in RNA, it would be extremely useful to identify small-molecule stand-ins for magnesium binding sites characterized by two or more inner-sphere coordinations[11]. Such ligands would provide useful probes for deciphering the energetic contributions of different types of ion binding sites in RNA structural stability, catalysis and ribonucleoprotein function.

Molecular mechanisms of ribozymes

Some of the most exciting recent advances in understanding the chemical mechanisms of ribozyme-catalyzed reactions have involved small-molecule inhibitors and activators of RNA active sites. In an approach that took its cue from a classic experiment with crystals of ribonuclease A, vanadate was used as a transition state analog in the catalytic center of the hairpin ribozyme[12]. Although the vanadate geometry is that of a distorted trigonal bipyramid, it is pentacoordinate and an approximate mimic of the transition state for phosphoryl transfer catalyzed by the ribozyme that would be otherwise impossible to visualize experimentally. The crystal structure of the hairpin ribozyme–vanadate complex, when compared with structures of the starting and product states, revealed a rigid active site that makes more hydrogen bonds to the transition state than to the precursor or product[12]. This finding corroborates the expectation that in this, and perhaps other, ribozymes, transition state stabilization is an important catalytic strategy.

In the hepatitis delta virus (HDV) self-cleaving ribozyme, where general acid-base catalysis was unexpectedly suggested by an initial crystal structure[13], imidazole and its chemical cousins provided important clues to the role of an active site cytidine in accelerating the rate of site-specific phosphodiester bond scission. Using a series of imidazole analogs with different pK_a values, it was possible to 'rescue' ribozymes rendered inactive by mutation of the critical active site cytidine to other nucleotides[14,15]. Correlation of reaction pK_as with those of various imidazole analogs supported a direct role of the critical cytidine in proton transfer during catalysis. Furthermore, sulfur substitution of the leaving-group oxygen, as well as 2′-*O*-methyl substitution of the attacking 2′-hydroxyl nucleophile, have enabled detailed kinetic and crystallographic analysis, respectively, of different HDV-ribozyme reaction states[16,17]. Although some details of the reaction mechanism remain unclear, the use of small-molecule activators and derivatives of the ribozyme have confirmed a general acid-base mechanism for this catalyst. RNA thus appears surprisingly capable of using its nucleobases, which have pK_as far from neutral, for proton shuttling analogous to that carried out within many protein active sites.

In some cases, single-atom changes within the active site of a large RNA have been sufficient to trap the RNA in a reaction state suitable for crystallographic analysis. For example, substitution of the terminal 2′-OH with 2′-H in the substrate of a group I intron



Suphoj Chancheaw

Figure 3 Approach to determining physiological activities of kinases by specific sensitization to chemical inhibitors. The ATP-binding site of a particular kinase is genetically modified to enable binding of an inhibitory analog that is too bulky to bind to native kinases. Substrates of the modified kinase can then be detected by identifying proteins, shown as bands in a polyacrylamide gel, that are no longer phosphorylated in cells expressing the mutated kinase in the presence of the inhibitor.

slowed intron-catalyzed splicing by ~1,000-fold, enabling crystallization of a catalytically active but stalled complex[18]. The resulting structure revealed two metal ions positioned 3.9 Å apart that were directly coordinated by the six biochemically predicted ligands. This observation provides compelling evidence that the two-metal-ion catalytic mechanism, common among proteins that catalyze phosphoryl transfer, is shared by this class of RNA catalysts.

Small molecules and small-molecule derivatized RNA will continue to have central roles in elucidating molecular structure and dynamics. For example, it would be exciting to use photoreactive nucleotide analogs to prevent ribozyme reaction in the crystalline state until activation by a laser[19]. Such an approach might enable trapping and visualization of transient reaction intermediates in a crystal. It will also be important, and is perhaps now feasible based on new chemical synthetic methods and availability of high-resolution structures, to develop libraries of specific chemical inhibitors of ribozymes and ribonucleoproteins that can be used both as basic research tools and as therapeutics. Here, the ribosome provides a paradigm as a large ribonucleoprotein and ribozyme that is a target of both naturally occurring and designed antimicrobials (**Fig. 2**; reviewed in ref. 20).

Biochemical and structural insights have begun to inform many of the new frontiers in RNA biology, including riboswitches, RNA interference and RNA-controlled translation initiation. Complementary chemical, structural and genetic approaches will now be essential for figuring out the molecular intricacies of these processes.

Physiological roles of molecular switches

Molecular structures provide chemical details that offer insight into biological mechanisms, but how do such findings relate to the physiological activities of these macromolecules in cells? Direct investigation of RNA and RNP complexes *in vivo* is in its infancy, but for other kinds of macromolecules, small molecules are already providing an essential tool for connecting chemical, structural and cell biology. One wonderful example of this has been the development of chemical tools for specifically inhibiting individual kinase enzymes *in vivo* using a combination of clever genetic manipulation and chemical insight. To do this, a specific amino acid in the ATP-binding site of a kinase is mutated to a small amino acid—glycine or alanine—thereby creating space for binding of a bulky ATP analog[21]. The analog simply does not fit in the active sites of native kinases because of steric inhibition. Thus, the mutation sensitizes just the kinase of interest to the inhibitor, and because the inhibitor ATP analogs are cell-permeable molecules, they can be readily introduced into cells to determine the physiological consequence of blocking the activity of a particular kinase enzyme (**Fig. 3**).

This approach has proven highly effective for dissecting the biological roles of specific kinases in cellular signaling pathways that control such fundamental processes as—progression through the cell cycle and cytokinesis (reviewed in ref. 22). The small molecules are fast acting, and binding is kinetically reversible, thereby avoiding potential cellular compensation that might otherwise mask a relevant phenotype. Furthermore, because many of the inhibitors used are drug-like compounds that strike an appropriate balance between aqueous solubility and cell permeability, they are potentially useful as therapeutics.

Can selective inhibitor design, via protein engineering coupled with chemical synthesis to define a binding site uniquely suited for small-molecule interaction, work for other kinds of molecules? One class that comes to mind is the GTP-hydrolyzing enzymes, a superfamily of proteins that are involved in signal transduction pathways and function as molecular switches in a variety of venues. For example, GTPases have central roles in both ribosome function (reviewed in ref. 23) and assembly[24–26] by advancing a complex of macromolecules through a series of ordered steps in response to cellular cues. Nucleotide analog inhibitors that inhibit GTPases would

therefore be extremely useful for examining the processes controlled by these ubiquitous regulators *in vivo*, as well as for trapping and determining structures of the large assemblies in which they participate.

Conclusions

Chemical and structural methods will go hand in hand into the future of macromolecular structural biology on all levels. Determining structures of large molecular assemblies by methods including crystallography, electron microscopy and other imaging techniques will require clever chemical means of trapping and tagging samples. Likewise, understanding the mechanistic roles of natural and designed small-molecule effectors in everything from individual enzymes to cell signaling and whole organism developmental pathways will necessitate structural and biochemical study. Perhaps most exciting is the increasing degree to which each discipline informs the other: clever chemical approaches make it possible to address new questions about biological function, from single molecules to entire systems, whereas burning biological problems drive the development of new chemical and biophysical methods. From its origins in small-molecule chemistry and biochemistry, chemical biology will offer many of the tools and insights necessary to answer outstanding questions in biology.

ACKNOWLEDGMENTS

I thank K. Berry and J. Cate for helpful discussions and comments on the manuscript.

1. Schreiber, S.L. *Nat. Chem. Biol.* **1**, 64–66 (2005).
2. Lay, P.A., Magnuson, R.H. & Taube, H. *Inorg. Chem.* **28**, 3001–3007 (1989).
3. Cate, J.H. & Doudna, J.A. *Structure* **4**, 1221–1229 (1996).
4. Cate, J.H. *et al. Science* **273**, 1678–1685 (1996).
5. Ban, N., Nissen, P., Hansen, J., Moore, P.B. & Steitz, T.A. *Science* **289**, 905–920 (2000).
6. Wimberly, B.T. *et al. Nature* **407**, 327–339 (2000).
7. Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N. & Noller, H.F. *Science* **285**, 2095–2104 (1999).
8. Batey, R.T., Gilbert, S.D. & Montange, R.K. *Nature* **432**, 411–415 (2004).
9. Kazantsev, A.V. *et al. Proc. Natl. Acad. Sci. USA* **102**, 13392–13397 (2005).
10. Torres-Larios, A., Swinger, K.K., Krasilnikov, A.S., Pan, T. & Mondragon, A. *Nature* **437**, 584–587 (2005).
11. Draper, D.E., Grilley, D. & Soto, A.M. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 221–243 (2005).
12. Rupert, P.B., Massey, A.P., Sigurdsson, S.T. & Ferre-D'Amare, A.R. *Science* **298**, 1421–1424 (2002).
13. Ferre-D'Amare, A.R., Zhou, K. & Doudna, J.A. *Nature* **395**, 567–574 (1998).
14. Perrotta, A.T., Shih, I. & Been, M.D. *Science* **286**, 123–126 (1999).
15. Shih, I.H. & Been, M.D. *Proc. Natl. Acad. Sci. USA* **98**, 1489–1494 (2001).
16. Das, S.R. & Piccirilli, J.A. *Nat. Chem. Biol.* **1**, 45–52 (2005).
17. Ke, A., Zhou, K., Ding, F., Cate, J.H. & Doudna, J.A. *Nature* **429**, 201–205 (2004).
18. Stahley, M.R. & Strobel, S.A. *Science* **309**, 1587–1590 (2005).
19. Chaulk, S.G. & MacMillan, A.M. *Nucleic Acids Res.* **26**, 3173–3178 (1998).
20. Sutcliffe, J.A. *Curr. Opin. Microbiol.* **8** 534–542 (2005).
21. Bishop, A.C. *et al. Nature* **407**, 395–401 (2000).
22. Bishop, A.C., Buzko, O. & Shokat, K.M. *Trends Cell Biol.* **11**, 167–172 (2001).
23. Nilsson, J. & Nissen, P. *Curr. Opin. Struct. Biol.* **15**, 349–354 (2005).
24. Gelperin, D., Horton, L., Beckman, J., Hensold, J. & Lemmon, S.K. *RNA* **7**, 1268–1283 (2001).
25. Wegierski, T., Billy, E., Nasr, F. & Filipowicz, W. *RNA* **7**, 1254–1267 (2001).
26. Karbstein, K. & Doudna, J.A. *Mol. Cell* (in the press).