Integrated multidisciplinarity in the natural sciences

Angela M. Gronenborn

From the Department of Structural Biology and Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260

Running title: Integrated multidisciplinarity in the natural sciences

*To whom correspondence should be addressed: Angela M. Gronenborn, Department of Structural Biology, University of Pittsburgh School of Medicine, Biomedical Science Tower 3, Room 1051, 3501 Fifth Avenue, Pittsburgh, PA 15260; E-mail: amg100@pitt.edu Telephone: (412) 648-9959; FAX: (412) 648-9008

Keywords: Discipline, natural sciences, integration, NMR, X-ray, cryoEM, structural biology, HIV-1 capsid, high-resolution structural model, multidisciplinary research

ABSTRACT

The integration of multiple perspectives in both the arts and natural sciences is tremendously powerful and arguably necessary for capturing relevant features of complex phenomena. Individual methods and models comprise abstractions from and idealizations of nature, and only the integration of multiple models, methods, and representations provides a means to reach more accurate results than relying on any single approach. In my Mildred Cohn Award Lecture at the 2019 ASBMB meeting, I illustrated the power of such multidisciplinary work by highlighting the successful integration of data and multiple views afforded by NMR spectroscopy, cryo-electron microscopy (cryo-EM), cryo-electron tomography (cryoET), X-ray crystallography, computation, and functional assays made possible through the collaborative efforts by members of the Pittsburgh Center for HIV Protein Interactions. This approach permitted us to generate the first all-atom model of a native HIV-1 capsid core.

For scholars in the arts, it comes as no surprise that a single object can elicit multiple, unique representations. In the case of the river Vltava (or 'The Moldau' in English), these different representations may take the form of photographs and maps, paintings by Egon Schiele, a poem by Bertolt Brecht, or the symphonic poem by Bedřich Smetana (Figure 1).

Smetana, in particular, was acutely aware of the aesthetic power inherent to multiple perspectives; he deliberately used the term 'tone painting’, to allude to the parallels between the visual and auditory domains. His creation Má vlast is a set of six pieces, composed between 1874 and 1879, one of which, Vltava, also known by its German name 'The Moldau', depicts the course of the river, following its path from two tiny trickles of water at the source in the mountains until it becomes a grand and majestic stream, flowing through Prague. Smetana's music reflects the landscapes, history and folklore of his native Bohemia and the quest of his people for independent statehood and freedom from German culture.

Another perspective is afforded in a series of paintings by Egon Schiele. Schiele retreated to Krumau at the Moldau, the birthplace of his mother, to escape from a perceived oppressive Viennese society in 1910. There, he sought to explore the spiritual essence of his environment, and his paintings hauntingly depict the claustrophobic nature of the ancient winding streets and the compact medieval townscape along the Moldau river, expressing his own melancholy reflections. Unfortunately, Schiele's stay in Krumau was cut short, the residents driving him out of town, being incensed by his bohemian appearance and unorthodox lifestyle.

The poem “The song of the Moldau” provides yet another distinctive vision of the river. The poem’s author, Bertolt Brecht, was one of the most influential writers, playwrights, and poets of the 20th century. In 1933, he fled Germany by way of Denmark, Finland and Russia to the United States but returned to Europe in 1949, after having been blacklisted in the US. Brecht had a supreme gift for
language. He wanted his audience to think and retain a critical detachment, in contrast to being solely captivated by emotions. He uses the invisible movements of stones at the bottom of the river as a metaphor for the inevitable changes in political power.

All three artists embraced a striking palette of colors in their works, reflecting a powerful obsession with a world torn apart by upheaval and depression. Each displayed a nonconformist, rebellious spirit in their work, and the 'Verfremdungseffekt', or alienation, can be seen in Schiele's paintings as well as Brecht's writings.

As illustrated by the work of the above artists, the depiction of an object depends on how it is viewed. The notion that identical entities can be characterized using different perspectives also holds true for the natural sciences. Representations of a single phenomenon can take on quite different forms, depending on whether the phenomenon is interrogated and described by a physicist, a chemist or a biologist. Therefore, it seems prudent to integrate these different views in order to achieve a more complete description of the phenomenon; Or, to quote Nietzsche: “Es gibt nur ein perspektivisches Sehen, nur ein perspektivisches ‘Erkennen’; und je mehr Affekte wir über eine Sache zu Worte kommen lassen, je mehr Augen, verschiedene Augen wir uns für dieselbe Sache einzusetzen wissen, um so vollständiger wird unser ‘Begriff’ dieser Sache, unsre ‘Objektivität’ sein.” (There only is a perspectival view, only perspectival recognition and the more pieces we allow to inform about an object, the more eyes, various eyes we are able to use for the same object, the more complete will be our ‘concept’ of an entity, our ‘objectivity’) (1). It is through the integration of different perspectives that we may succeed to comprehend the world around us, or, as Dr. Faustus states, understand “was die Welt im Innersten zusammenhält” (whatever binds the world’s innermost core together) (2).

There are clear parallels to the above logic in the natural sciences. The accumulation of more and more scientific knowledge invariably results in specialization because of divisions in ideology, epistemology, methodology or theory. As a result, any individual discipline becomes ‘unknowable’ - no single theory nor conceptual framework can continue to encompass the entire field. This causes fragmentation and, when the fragments recombine, chimera, or new disciplines, appear (Figure 2). Simultaneously, disciplinary boundaries are constantly recreated and reshaped by the research enterprise. In fact, more and more, scientists from diverse disciplines come together and work collectively, combining their knowledge and approaches to find answers to increasingly complex scientific questions. This represents a shift from the early days of science, which were the exclusive province of the lone researcher.

At the present time, the most interesting problems in science, and many of the most important facing society, require research at the interfaces between traditional disciplines. Examples include: understanding life as coordinated networks of chemical reactions; production, storage and conservation of energy and water; the management of greenhouse gases and elucidating the molecular basis of disease. As a consequence, more and more research will need to be conducted at the interfaces of the once “traditional” disciplines. Such interfaces exist between biology and chemistry, biology and physics, mathematics and biology, pharmacology and chemistry, engineering and biology, material science and biology and many more (Figure 2). In order to find the most appropriate solutions to any of the above listed complex problems, we need to integrate multiple approaches and the resultant findings across the disciplines.

Indeed, a common notion today is that interdisciplinary research leads to breakthroughs in science – the theme of high risk-high reward is frequently evoked and well supported by anecdotal evidence; just think of the landmark papers on the DNA structure (3-5). It was through the combined efforts of Rosalind Franklin, a chemist, Maurice Wilkins, a physicist, Jim Watson, a biologist, and Francis Crick, a physicist, that this breakthrough discovery was achieved (6).

In my own research arena, I posit that the work emerging from the interdisciplinary Pittsburgh Center for HIV Protein Interactions (PCHPI; http://www.hivppi.pitt.edu/), represents an illustrative example of integrative multidisciplinary team science. Within the center, virologists, cell biologists and structural biologists are engaged in a collaborative effort to push the field of HIV biology forward with the hope that the knowledge gained can be effectively leveraged for developing novel therapeutic strategies.
The primary emphasis of the PCHPI research program is to define events, pathways, and host cell factors that are important in the HIV replication cycle, particularly those during the early stages of infection. Virtually every step in HIV replication involves an intricate interplay between the virus and the host machinery, and cellular components play pivotal roles during viral entry, reverse transcription, nuclear import, integration, transcription, nuclear export, translation, assembly, and budding. Indeed, it has become apparent that the virus utilizes the host machinery both to promote its replication and, at the same time, to subvert and evade the antiviral responses of the cell. PCHPI members are dedicated to structurally characterizing HIV proteins and their complexes with host protein interacting partners by X-ray crystallography, solution and solid state nuclear magnetic resonance spectroscopy (NMR), cryo-electron microscopy (cryo-EM) and cellular imaging methodologies. Naturally, all the structural techniques constitute a smooth continuum of methodologies and, for a complete and comprehensive picture to emerge, they all have an important role to play. In addition, structural methodologies need to be integrated with cell biology, biochemistry and a host of biological investigations in order to interpret structures with respect to function - clearly an important aim and reason for structural work. In addition, computation is playing an increasingly critical role, especially for very large and complex structures and molecular machines.

Below, I illustrate the integrative structural biology work performed by the PCHPI using the HIV capsid core structure as an example (Figure 3).

Mature HIV-1 virions contain a conical-shaped shell (core), enclosing the viral RNA genome. The native core structure is essential for successful viral replication. It comprises ~1500 copies of the capsid protein (CA) with the CA polypeptide chain folded into two domains, a ~150 amino acid N-terminal domain (NTD) and a ~80 amino acid C-terminal domain (CTD), connected by a hinge. While the NTD is monomeric in solution, the CTD dimerizes with a $K_d$ of 10µM, and dimerization is dependent on the presence of residues Trp184 and Met185 (7). Dimerization via the CTDs is also at play during assembly of the full-length protein into the mature capsid core; mutagenesis of Trp184 and Met185 interferes with in vitro assembly and abolishes viral infectivity in cell culture (8). When assembled in vitro, full-length CA forms single-walled tube structures, and early cryo-EM reconstructions of these assemblies revealed a hexameric arrangement of the NTDs on the outside of the tubes (9,10), with three helices from each chain combining with those from adjacent chains to form an 18-helix bundle at the center of a hexameric unit. In the overall lattice, neighboring hexamers are connected via the dimeric CTD. In terms of dimensions, assembled CA tubes have a wall thickness of ~80 Å, corresponding to the height of the CA protein, and an overall tube diameter of ~420 Å, similar to the diameters measured for native conical cores, which range from ~440 Å to ~590 Å (11).

We determined the solution NMR structure of the dimeric CTD (12) and found that it differed from several crystal structures of CTD dimers (7,13,14). While the structures of the single CTD chains are all very similar, considerable variability at the dimerization interface was noted, primarily involving different crossing angle arrangements of the two antiparallel helices that form the dimer contacts. To complement our NMR data and assess whether the dimerization interface that is present in solution is also present in assembled CA structures, cryo-EM data were collected for tubes of different helical symmetries. An initial cryo-EM model was generated starting for tubes that belonged to the (-13,11) helical family. Consistent with the previous findings (9), the NTDs form hexameric rings on the outer tube surface, while the CTDs line the inner tube surface. In order to generate a pseudo-atomic model, the X-ray structure of the NTD (PDB: 2GON) (15) and the solution NMR structure of the CTD dimer (PDB: 2KOD (12)) were independently fitted into the tubular density map, and an excellent fit was obtained. Interestingly, in addition to density around the pseudo six-fold axis and the two-fold axis at the CTD dimer interface, we also observed density at a three-fold axis in the non-symmetrized density map, at the center of three CTD dimers. Reconstruction of a pseudo-atomic model of three adjacent CA hexamers revealed that three pairs of helices, H10 from one CTD and H11 from a neighboring CTD, constitute this interface. The importance of this interface for viral infectivity and integrity of the core was probed by mutagenesis. Functional studies with mutations at the three-fold interface confirmed its importance for optimal viral...
activity. Intriguingly, stabilizing the interface by Cys-Cys crosslinking reduced infectivity, possibly by interfering with the correct timing of uncoating or by interfering with the plasticity that is required for forming the various types of curvatures inherent to native, asymmetric retroviral capsid cores.

To derive a model of a conical HIV capsid core, an atomic model for a tubular assembly had to be constructed. To that end, atomic structures of the NTD (PDB: 2GON (15) or 3H47 (16)) and CTD (PDB: 2KOD (12)) were docked into the cryo-EM density map, linker and loop regions were generated using homology modelling, and molecular dynamics flexible fitting (MDFF) was applied (17). In this manner, a portion of an in vitro assembled CA tube was generated, comprising 71 CA hexamers (13 million atoms, including water molecules and ions) (18).

The curvature needed to assemble a conical core from hexameric building blocks is imparted by plasticity in the CTD, primarily around the helix H9 dimer interface. However, as was already known in the 15th century and illustrated by Leonardo da Vinci (19), closure of an ovoid is impossible using a hexameric lattice, and twelve pentameric units are required. To generate the building blocks of the core, a hexamer-of-hexamers (HOH) patch (six hexamers surrounding a central hexamer) was extracted from the helical tube assembly, and a model for a pentamer of hexamers (POH; five hexamers surrounding a central pentamer) was created by replacing the central hexamer with a pentamer. The latter used as the starting model the X-ray structure of a mutant capsid protein crosslinked into a pentameric unit (PDB: 3P05) (20). Although the initial POH contained some gaps between the surrounding hexamers, MD simulations permitted closure of these spaces, and the predominantly flat starting model transformed into a highly curved, dome-like structure, with no steric clashes or breaks at the hexamer seams (18). Thus, the essential building blocks for the generation a realistic atomic conical capsid core structure were in hand.

In order to generate a realistic all atom model of a native capsid, it was necessary to isolate and purify cores from virions, given the pleiomorphic nature of native HIV cores. To this end, cryo-electron tomography (cryo-ET) was performed and a series of tomographic slices for a common, high-quality HIV-1 core with a cone angle of \( \sim 23^\circ \) was collected. Guided by the shape, size and structural features of the capsid layer, several classes of theoretical fullerene models were generated, including classes with 252, 216, 186 or 166 hexamers. These were evaluated and tested by cross-correlation between the model and the capsid density and two models were used to generate all-atom molecular dynamics models of an entire HIV-1 capsid, either possessing 1,356 or 1,176 CA chains. The final best model contained 216 hexamer units and 12 pentamer units (1356 CA chains). This assembly was subjected to an unconstrained 100 ns molecular dynamics simulation, in a box of water, under physiological salt conditions, thus comprising, in total, 64 million atoms (18). This all-atom HIV-1 capsid core was stable throughout the entire length of the simulation.

As illustrated by this example from the PCHPI, bringing together groups of scientists with varied disciplinary backgrounds and training (physics, chemistry, virology) allowed us to tackle a complex biological system and to achieve its structural characterization. By applying an integrated structural and computational approach, we now have in hand an all-atom HIV capsid core model, which serves as a valuable platform for further experiments of capsid function and for targeted pharmacological intervention.

At this point, let me close with some philosophical thoughts. Complex biological phenomena are characterized by numerous features, all involved in causal behavior. They frequently act at various scales or levels of organization. As a result, there is a need to employ multiple methods, models, and representations to promote scientific understanding. I contend that the relationship among multiple models needs to be one of integrative pluralism. This is necessitated by the fact that, in some models, certain attributes or features are necessarily relegated to the undescribed context whereas, in other models, these same features are explicitly represented. Thus, any one methodological approach alone results in partiality of representation. For example, NMR-derived models infer atomic positions from magnetization transfer properties between nuclei, while in crystallography or single particle cryo-EM, atomic positions are obtained from scattering of the X-ray or electron beam, respectively. As a result of the inherent partiality, many compatible models are
necessary to reach as accurate a description as possible.

So, why do I emphasize the above in an article based on my 2019 Mildred Cohn Award Lecture at the ASBMB meeting? Mildred Cohn understood and practiced what we would now call integrated multidisciplinarity in her research. She started out by establishing the use of stable isotopic tracers to study biochemical reactions, such as following the metabolism of sulfur-containing compounds (21). She used mass-spectrometry as a primary technique in her early research, building her own instruments for this purpose. In the 1950s, at Washington University, her interests focused on reactions using ATP, both with respect to the divalent metal ion involvement as well as the specificity of bond cleavage. She first used ESR spectroscopy to get answers (22) and, after this methodology failed her, embarked on NMR, in particular $^{31}$P NMR. However, it took several years before she managed to persuade Varian to let her run $^{31}$P spectra, with which she demonstrated the feasibility of distinguishing the $\alpha$-, $\beta$- and $\gamma$-phosphorous resonances, causing her to embrace NMR. After moving to the Johnson Research Foundation at the University of Pennsylvania, Cohn combined ESR, EPR and NMR approaches with meticulous biochemistry to study enzyme mechanisms (23). She clearly was not wetted to a single approach or methodology; she integrated the findings from multiple methods and in this manner succeeded to make pioneering and important contributions to enzymology.

Acknowledgements: I thank all former and present members of the PCHPI and the Gronenborn laboratory for their contributions to the studies referenced here and Teresa Brosenitsch for excellent editorial help. I also want to express special thanks to Sandra Mitchell, Department of History and Philosophy of Science, University of Pittsburgh, with whom, over the past several years, I have had frequent conversations about perspectives, science and more, and who prompts me to engage different sets of neurons of my brain that otherwise would sit idle.

Conflict of Interest: The author declares no conflict of interest.

Supporting Information: English translation of the Bertolt Brecht poem of Figure 1.
References

1. Nietzsche, F. (1967) Critical edition of the complete works and letters, based on the critical text by G. Colli and M. Montinari, de Gruyter, Berlin/New York.
2. von Goethe, J. W. (2018) Faust: Der Tragodie Erster Theil, Szene Nacht. in Faust edition komplett: Gesamthandschrift und konstituierter Text (Bohnenkamp, A., Henke, S., and Jannidi, F. eds.), Wallstein-Verlag, Gottingen.
3. Watson, J. D., and Crick, F. H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738
4. Wilkins, M. H., Stokes, A. R., and Wilson, H. R. (1953) Molecular structure of deoxypentose nucleic acids. *Nature* **171**, 738-740
5. Franklin, R. E., and Gosling, R. G. (1953) Molecular configuration in sodium thymonucleate. *Nature* **171**, 740-741
6. Maddox, B. (2002) Rosalind Franklin: the dark lady of DNA, 1st ed., Harper Collins, New York.
7. Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worhyleake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **278**, 849-853
8. von Schwedler, U. K., Stray, K. M., Garrus, J. E., and Sundquist, W. I. (2003) Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J Virol* **77**, 5439-5450
9. Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409-413
10. Ganser-Pornillos, B. K., Cheng, A., and Yeager, M. (2007) Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell* **131**, 70-79
11. Briggs, J. A., Wilk, T., Welker, R., Krausslich, H. G., and Fuller, S. D. (2003) Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J* **22**, 1707-1715
12. Byeon, I. J., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., Shi, J., Concel, J., Aiken, C., Zhang, P., and Gronenborn, A. M. (2009) Structural convergence between cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* **139**, 780-790
13. Ternois, F., Sticht, J., Duquerroy, S., Krausslich, H. G., and Rey, F. A. (2005) The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat Struct Mol Biol* **12**, 678-682
14. Worhyleake, D. K., Wang, H., Yoo, S., Sundquist, W. I., and Hill, C. P. (1999) Structures of the HIV-1 capsid protein dimerization domain at 2.6 A resolution. *Acta Crystallogr D Biol Crystallogr* **55**, 85-92
15. Kelly, B. N., Howard, B. R., Wang, H., Robinson, H., Sundquist, W. I., and Hill, C. P. (2006) Implications for viral capsid assembly from crystal structures of HIV-1 Gag (1-278) and CA(N)(133-278). *Biochemistry* **45**, 11257-11266
16. Pornillos, O., Ganser-Pornillos, B. K., Kelly, B. N., Hua, Y., Whitby, F. G., Stout, C. D., Sundquist, W. I., Hill, C. P., and Yeager, M. (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**, 1282-1292
17. Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B., and Schulten, K. (2009) Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods* **49**, 174-180
18. Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A. M., Schulten, K., Aiken, C., and Zhang, P. (2013) Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643-646
19. Pacioli, L. (1509) *De divina proportione* (English: *On the Divine Proportion*), illustrated by Leonardo da Vinci; Antonio Capella, Paganino Paganini, Venice
20. Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011) Atomic-level modelling of the HIV capsid. *Nature* **469**, 424-427
21. du Vigneaud, V., Cohn, M., Chandler, J. P., Schenk, J. R., and Simmonds, S. (1941) The utilization of the methyl group of methionine in the biological synthesis of choline and creatine. *J Biol Chem* **140**, 625-641

22. Cohn, M., and Townsend, J. (1954) A study of manganous complexes by paramagnetic resonance adsorption. *Nature* **173**, 1090

23. Cohn, M., and Reed, G. H. (1982) Magnetic resonance studies of active sites in enzymic complexes. *Annu Rev Biochem* **51**, 365-394
Footnotes: The work of the PCHPI was supported by the National Institutes of Health grant (P50 GM082251). The abbreviations used are: CA, capsid protein; CTD, C-terminal domain; cryo-EM, cryo-electron microscopy; cryo-ET, cryo-electron tomography; EPR, electron paramagnetic resonance ESR, electron spin resonance; MDFF, molecular dynamics flexible fitting; NMR, nuclear magnetic resonance; NTD, N-terminal domain; PCHPI, Pittsburgh Center for HIV Protein Interactions; Protein data bank, PDB.

Figure Legends:

Figure 1. Four representations of the Czech river Vltava (Moldau). A photograph of Krumau at the Moldau (top left, with permission from Jürgen Reichmann), a painting of the same by Egon Schiele (bottom left), “The song of the Moldau” by Bertolt Brecht (top right, see supplemental information for English translation), and the first page of the Vltava score by Bedřich Smetana (bottom right).

Figure 2. Schematic representation of the relationship between traditional and hybrid disciplines. Hybrid disciplines arise by division and recombination of mature, traditional disciplines.

Figure 3. Integration of multiple methodologies at the Pittsburgh Center for HIV Protein Interactions. Data obtained by cryo-EM, NMR spectroscopy, X-ray crystallography, and cryo-ET were combined using computational methods and integrated with functional studies to derive the first all-atom model of a functional HIV-1 capsid. Data and images are from references (12) and (18).
Das Lied von der Moldau
Bertolt Brecht; 1943

Am Grunde der Moldau wandern die Steine
Es liegen drei Kaiser begraben in Prag.
Das Große bleibt groß nicht und klein nicht das Kleine.
Die Nacht hat zwölf Stunden, dann kommt schon der Tag.

Es wechseln die Zeiten. Die riesigen Pläne
Der Mächtigen kommen am Ende zum Halt.
Und gehn sie einher auch wie blutige Hähne
Es wechseln die Zeiten, da hilft kein Gewalt.

Am Grunde der Moldau wandern die Steine
Es liegen drei Kaiser begraben in Prag.
Das Große bleibt groß nicht und klein nicht das Kleine.
Die Nacht hat zwölf Stunden, dann kommt schon der Tag.
Fig. 2
Cryo-EM density map of capsid tubes

X-ray structure of NTD

NMR structure of CTD dimer

71 hexamers (13 million atoms)
10 ns MDFF into density

Mutational analysis of the interface in vivo

Fig. 3
