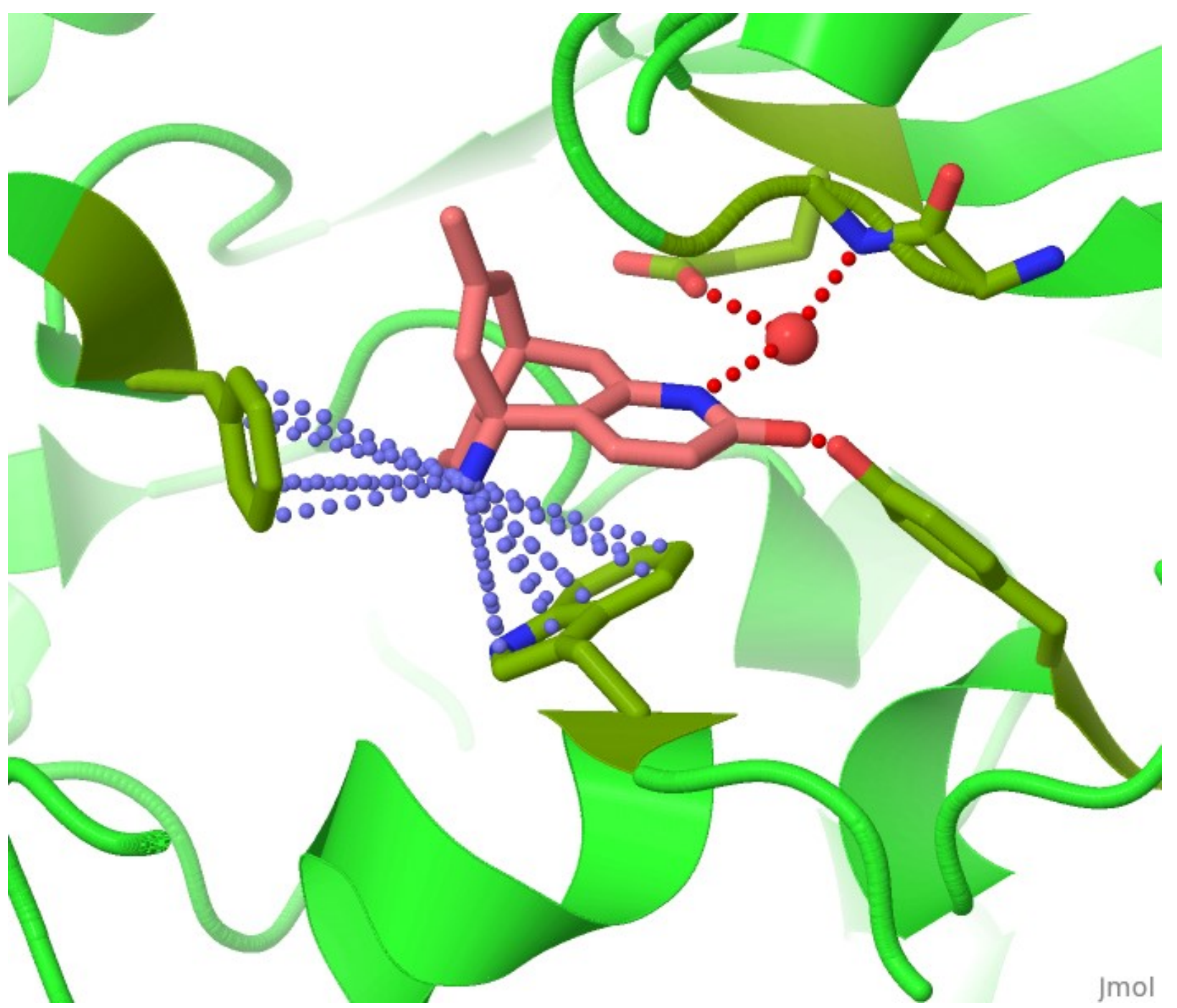# L36. Cloud-Based Visualization of Value-Added Model Annotations Using Jmol

Robert M. Hanson

St. Olaf College, Northfield, Minnesota
ISMB-ECCB 2015
Dublin, Ireland, July 10-14, 2015

Abstract

Recent advances in cloud-based services now allow the real-time merging of 3D structural models of proteins and nucleic acids from wwPDB with calculational "annotations" -- model validation information for deposited structures, secondary structure analysis for RNA and DNA, sequence alignment information for proteins. This poster presents recent developments in Jmol carried out in collaboration with the European Bioinformatics Institute (EBI) and others that change the meaning of "visualization" with respect to biomolecules, extending such visualization in ways that have not been possible previously.

The Problem

The PDB database, now with over 110,000 entries, is recognized as the premier repository for biomolecular structures. But it is important to remember that the reported "structures" of biomolecules are in many respects idealized models. The deposition process involves a variety of validation steps and yet, even after a structure is accepted in the PDB, many aspects of a structure may be unknown, known to be highly approximate, or simply missing.

Understanding the limitations of these published model structures is important if they are to be used responsibly. And yet, how is one to gain that understanding? Recent advancements at the wwPDB allow for a certain amount of investigation. The *Structure Validation Score Card* at RCSB, for example, gives hints in a subset of general categories, shown here for one for the most recently deposited structures, 4TVJ.
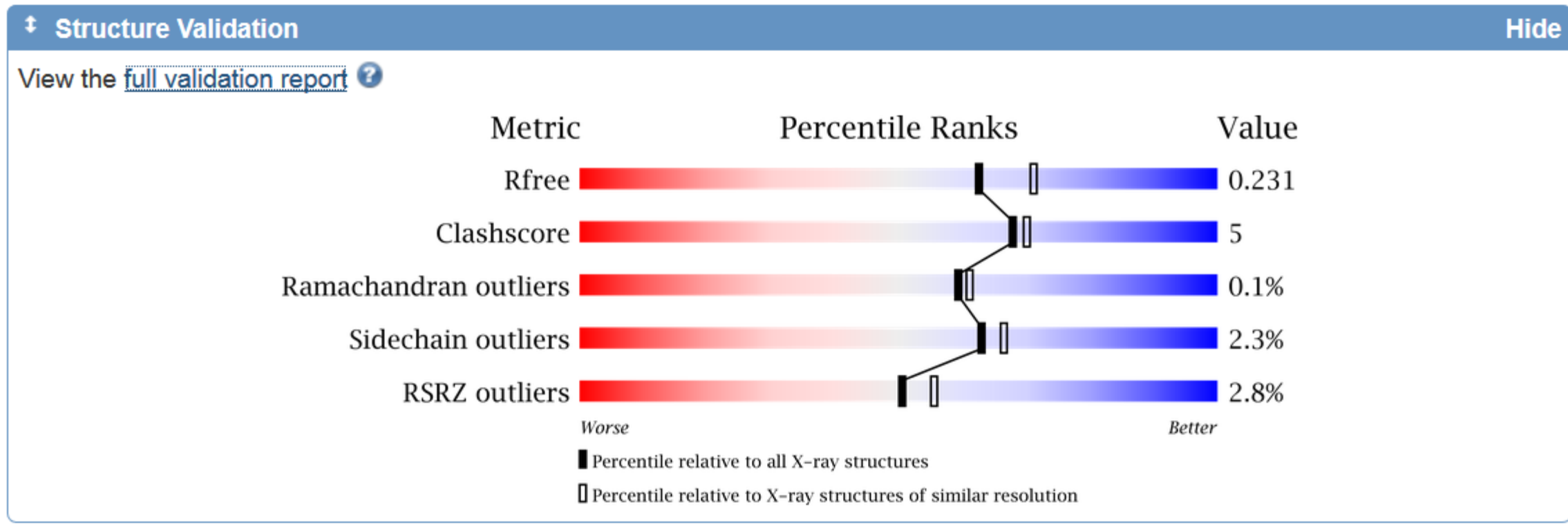
And a *full validation report* is also available, with its detailed tables.

Mol	Chain	Res	Type	RSRZ	Atom-1	Atom-2	Distance(Å)	Clash(Å)
1	B	226	THR	6.5	1:B:529:VAL:HG13	1:B:530:THR:H	1.81	0.46
1	A	547	LEU	4.5	1:A:377:TYR:CZ	1:A:381:VAL:HG21	2.51	0.46
1	A	549	PRO	4.5	1:A:472:ASN:HA	3:A:602:GOL:H2	1.98	0.45
1	A	224	LEU	4.2	1:A:355:GLU:OE2	1:A:363:ARG:NH1	2.40	0.45
1	B	354	PRO	4.1	1:B:323:THR:CG2	1:B:325:LYS:HG3	2.46	0.45
1	A	548	ASN	4.0	1:B:341:GLU:OE2	1:B:345:LYS:NZ	2.47	0.44
1	B	550	ASP	3.9	1:B:389:THR:HB	1:B:466:MET:HA	2.00	0.43
1	A	550	ASP	3.6	1:A:421:HIS:NE2	1:A:494:GLN:HG2	2.33	0.43
1	B	547	LEU	3.2	1:A:293:GLN:O	1:A:294:HIS:ND1	2.52	0.43
1	A	297	ALA	3.1	1:B:252:GLU:O	1:B:256:MET:HG3	2.19	0.43
1	B	549	PRO	2.9	1:A:570:ARG:NH1	4:A:740:HOH:O	2.52	0.42
1	A	223	ASP	2.9	1:B:531:LEU:HD12	1:B:531:LEU:HA	1.93	0.42
1	A	552	TYR	2.8				
1	B	349	THR	2.8				
1	B	297	ALA	2.7				
1	A	300	GLU	2.6				
1	B	226	THR	2.5				
1	A	429	GLY	2.2				
1	A	296	ARG	2.2				
1	B	296	ARG	2.0				

The Solution – Validation "Annotations" in Jmol

In collaboration with the European Bioinformatics Institute, PDB*e*, and the Research Consortium for Structural Biology (RCSB), we designed a method of accessing the validation data for all deposited PDB structures. Our design goals included:

(a) easy retrieval of structure-based annotation information,
(b) immediate access to visualization of that information, and
(c) utilization of the emerging value-added cloud resources.

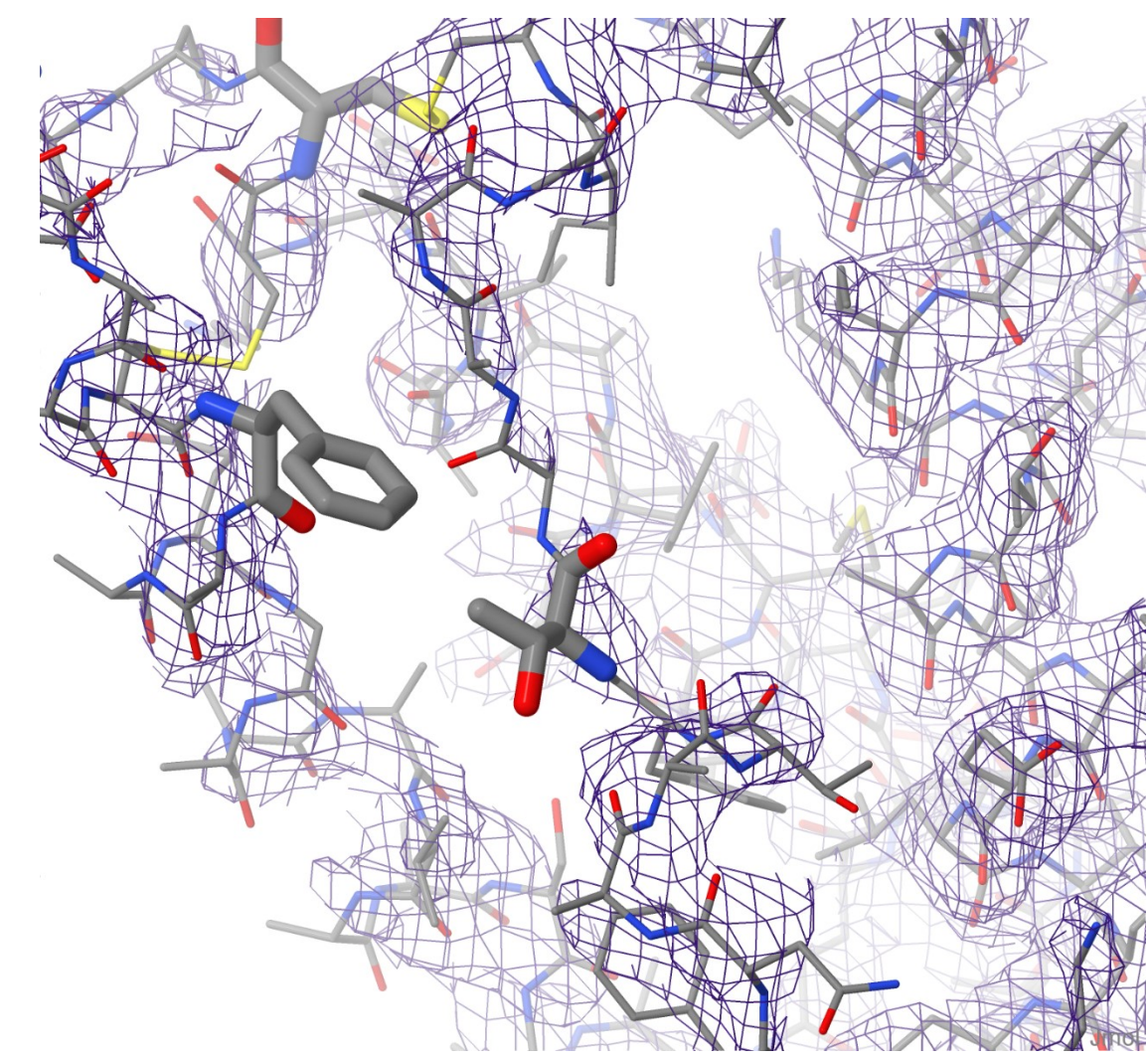In June of 2014 we embarked on a project to accomplish these goals.

EBI -- PDBe

St. Olaf -- Jmol

RCSB -- PDB
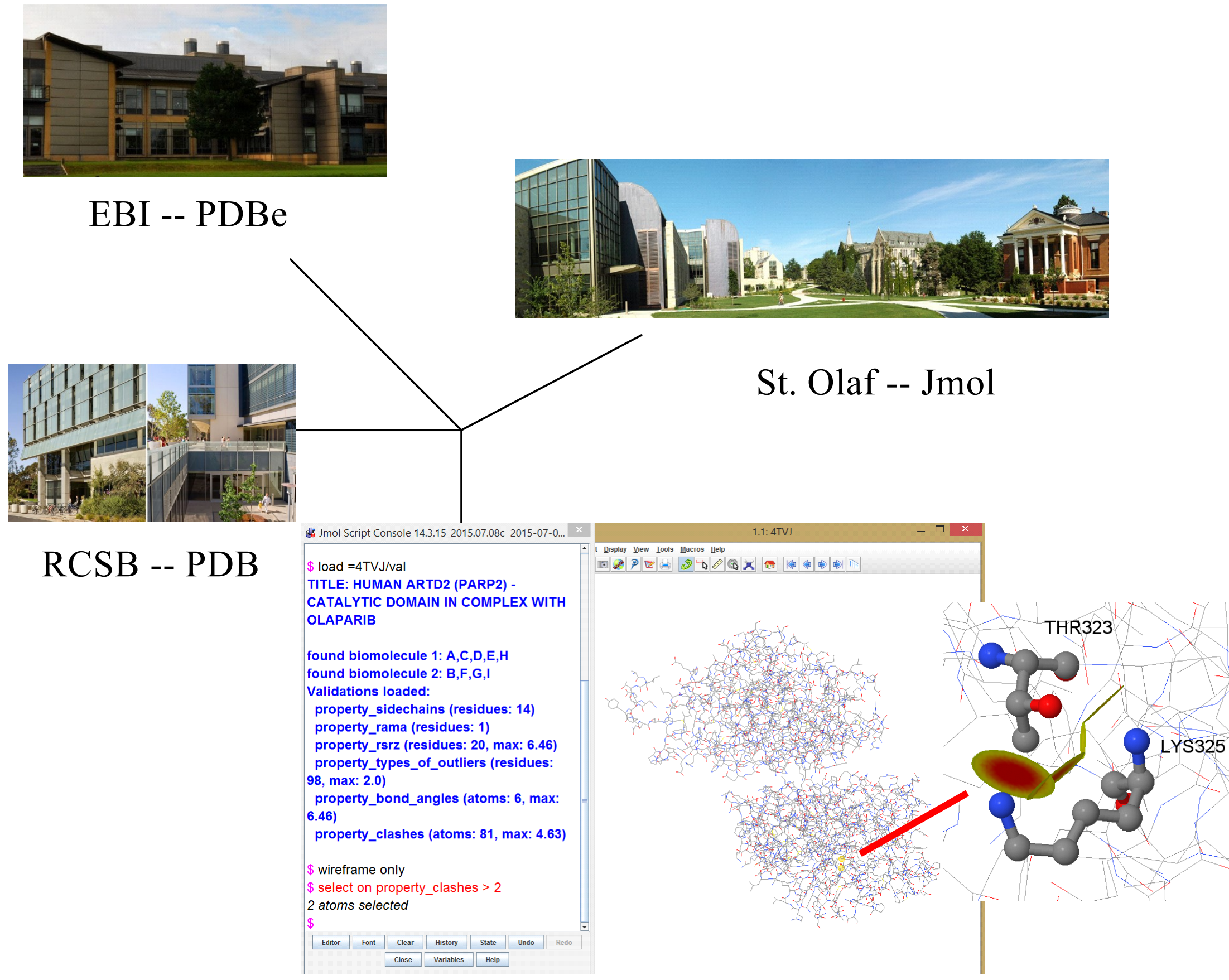
For several years it has been very easy to retrieve structures from RCSB (**load =4TVJ**) or from PDBe (**load *4TVJ**) into Jmol. What is new is the capability to add to that structural data the **entire set of validation data present in the validation report.** This is accomplished by appending to the Jmol **load** request the suffix */val*: **load =4TVJ/val**; **load *4TVJ/val**. The */val* suffix instructs Jmol to make an associated call to PDBe's *RESTful* API in order to retrieve a structured JSON list of validation data that then becomes part of the overall Jmol data structure associated with this model.
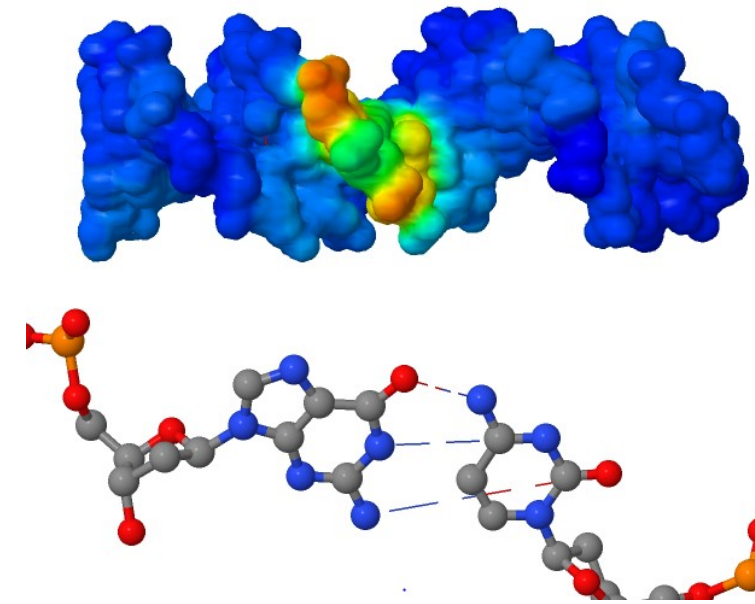
```
"4tvj": {
  "bond_angles": {
    "outliers": [
      {
        "units": [
          "|1|B|ARG|316|NE|||",
          "|1|B|ARG|316|CZ|||",
          "|1|B|ARG|316|NH2|||"
        ],
        "value": -6.46
      },
      {
        "units": [
          "|1|B|ARG|478|NE|||",
          "|1|B|ARG|478|CZ|||",
          "|1|B|ARG|478|NH1|||"
        ],
        "value": 5.41
      }
    ]
  },
  "chirals": {},
  "rsrz": {
    "outliers": [
      {
        "units": [
          "|1|A|ASP|223||||"
        ],
        "value": 2.891
```

Results

A variety of annotation visualizations in Jmol are shown below, either directly from EBI or also in association with electron density data from the Uppsala University Electron Density Server (http://eds.bmc.uu.se/eds) or the DSSR server at Columbia University (http://x3dna.org) using the annotation notation **load =1bna/dssr**

a section of GAL4 DNA complement (1d66) colored by straightness, identifying a base accidentally flipped during modeling and not noticed by the author or reviewers

load =1d66; set quaternionFrame "C"; isosurface select {dna} only vdw map property straightness
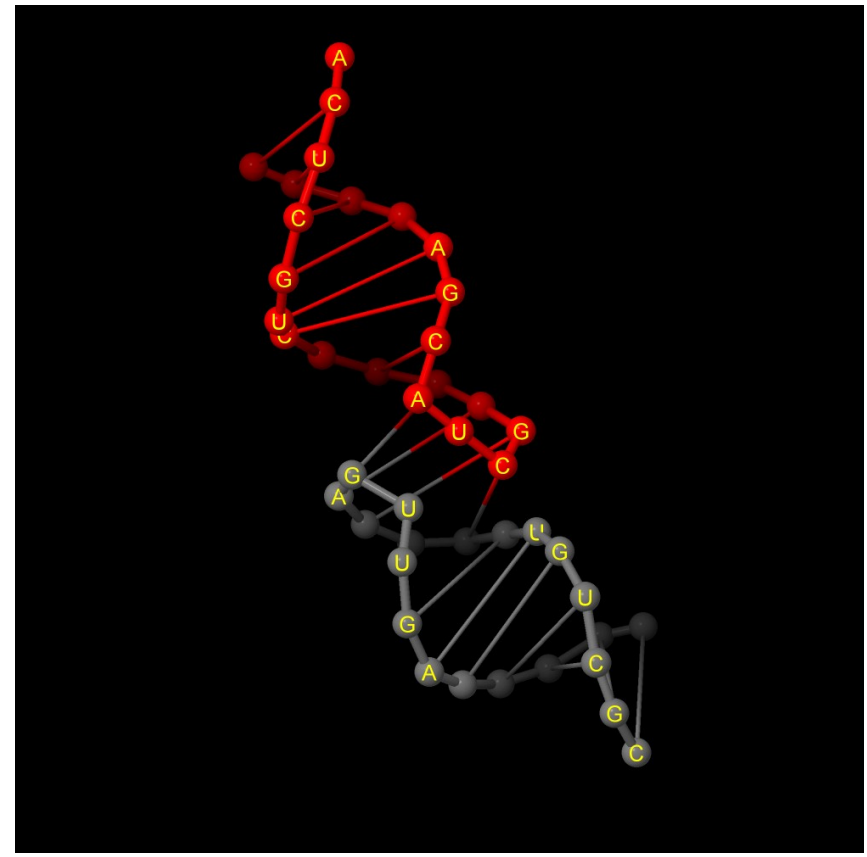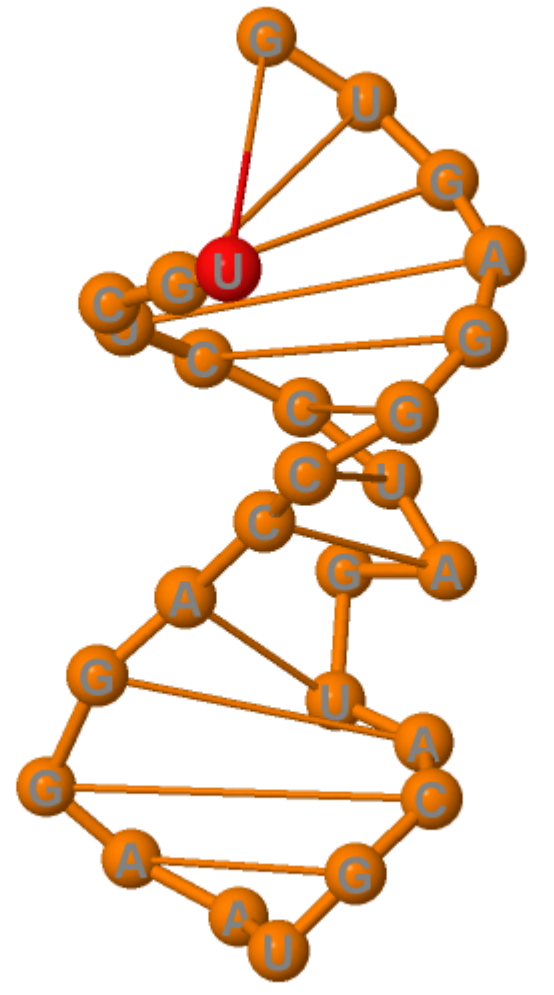
Missing electron density visualized using
load =2BXA/val; wireframe -0.1; select group property_rsrz > 5; wireframe 0.3; isosurface "=2BXA"; set zshade true

a kissing loop
load =2mi0/dssr; display within(dssr,"coaxialStacks")

the GUUA tetraloop in 1msy
load =1msy/dssr

Conclusions

Structural annotations, particularly when made easy to access, can significantly enhance molecular data. By associating these annotations with a program such as Jmol, which can load and visualize them easily, additional meaning can be derived, and valuable lessons can be learned about the meaning of a "biomolecular model" and its limitations.

Acknowledgments