Modeling Protein Structure of Ac Transposase
and other hAT Transpoases

Xianyan Kuang

Vollbrecht lab/Brendel lab
GDCB Dept/BCB Program
Ac/Ds transposon and Ac Transposase

- hAT superfamily.

- A classical two-component DNA transposon, and one of the best-characterized transposons in plants/eukaryotes.

- No protein structure resolved;

- One hAT member Hermes in insects has its structure resolved.

Specific project goals:

- Modeling protein structure of Ac transposase
  Evaluate modeled structures using prior functional knowledge
- Modeling protein structure of other hAT superfamily members
  Infer protein structure based phylogeny (initial effort; ongoing)
General procedure

![Diagram of the general procedure for protein structure prediction](http://www.pdg.cnb.uam.es/cursos/oeiras2005/practicass/StructPred.jpg)

1. **Experimental Sequence** → **Database Searching**
2. **Secondary Structure Prediction**
   - **No** → **Fold Prediction**
   - **Yes** → **Homology Modelling**
3. **Structural Homologue?**
   - **No**
   - **Yes** → **Final Structure??**

Ac Transposase: functional features defined by **genetic** and **biochemical** approaches

---

**Figure 1.** Autonomous plant hAT elements. The diagram below Ac shows the domains found within the Ac TPase protein. b, basic regions; x, N-terminal fragment that can be removed without loss of TPase function; m, middle region where the reactive residues are likely to occur; N1, N2 + 3, nuclear localization signals; PQ, Pro-Glx repeat region essential for transposition; DNA, DNA binding domain; DIM, dimerization domain: hAT1, hAT2, hAT3, regions of strong sequence similarity among hAT family TPase proteins, including those in animals and fungi. Introns are indicated by white bars within transcribed (hatched) region.

Craig et al, 2002 Mobile DNA II
However, the experimentally defined domains do not have structural homologues…

- Too short domains
- Sequence identity too low

So,,, secondary structure prediction…
DDE Recombinase = RNaseH like fold + “inserted domain”

HIV1 Integrase (3L3U)
5 sheets; 6 helixes

Ac/Ds (and hAT superfamily): Hairpin structure intermediate?!

Tn5 (1MUR)
Prokaryote transposase complex (with DNA); hairpin formation in transposon end
Comparison of recombinase sequences and predicted secondary structures

Herm: (141)  LFSITLRSKCVSDPAKEKALISRECIAVE---------KD---ASATSILWD---------  184
Hobo: (189)  LFDITLSRKAKSDPAEEKSSLIE11KGYD------S-G---ASATKWID---------  232
Ac: (263)    PIKSRTVACMYIMDYLEKEKCLKGDKY---Q---S-R---FSITAVWIS---------  305
Tam3: (297)  KISRA9CFRDQGYEKEH11RNLVEFHKG--------FNG---RISLIVAVWG---------  279
HIV IN: (50)  MNGQPV-DC-SFG---IWQIQGTCL---------  65
Tn5 Tnp: (67)  FNYSAALIKKAAGAMQTVKLQ----------E---FP---ELA9HLHSYQDAELGK  113
RAG1: (547)  EYFPDIJMMKFRNDSAVSLVMNDEDELGMRQSDLIDIVNGFTVVKESNGDG---YEKLQ---  610

6 predicted α-helix;
5 predicted β-strand.

Conserved DDE motif: boxed.

CxxH and C2H2 motifs: conserved among hAT elements and RAG-1 noted (∞).

The green box includes a region of Tn5 known to be involved in base flipping and hairpin formation and regions of hAT transposases and RAG1 presumably involved in hairpin formation. The basic amino acids marked with “*” indicate positions in RAG-1 where sequence changes resulted in a defect in hairpin formation.

(Hickman et al.; 2005 Nature Structural & Molecular Biology Suppl. Fig 5)
Fold prediction hybrid methods

- Swiss-Model server
- I-TASSER
Model selection/evaluation: Ac vs hermes

Ac (maize)

Hermes (house fly)
X-ray crystal structure (2bw3)

QH=0.7912; RMSD=1.0105, Percent Identity=16.23
Ac Transposase: the catalytic core compared with Hermes the template

Hermes x-ray crystal structure

(Hickman et al.; 2005 Nature Structural & Molecular Biology Fig 1)
Ac Transposase: the catalytic core compared with Hermes the template

Diagram showing structural comparisons of different domains and motifs, such as Retroviral integrase RNaseH-like fold and "inserted" domain.
Protein structure models of other hAT Transposases

Hermes (house fly)
X-ray crystal structure (2bw3)

+ Well-supported structure model

Ac (maize)

Protein structure based Phylogeny?
Inferring structure-based phylogeny:

- Collect protein sequences;
- Swiss-model server;
- Model selection: functional motif and structural alignment with template.

Structure and Evolution of the HAT transposon superfamily (Rubin et al, 2000 Genetics)
Superimposition of 8 hAT transposases in fungi, plants and animals

Catalytic core
(RNaseH fold and “inserted” domain)

Key motifs: DDE active site, CxxH zinc finger DNA binding motif and putative hairpin formation residue Trp (W).
The structure-based phylogeny, different from the sequence-based version, may better reflect the evolution of transposase functionality in hAT superfamily.
Conclusions

• Model protein structure of Ac transposase using threading method. 
well supported model by functional domain/motif assessment.

• Model protein structure of some other hAT transposases
conserved configuration at the entire structure level,
more strikingly conserved structure at the DDE catalytic site.
structure-based phylogeny differs from sequence-based phylogeny.

Future directions

• Model transposase/DNA complex?
  The CxxH zinc finger motif
  The Tn5 prokaryote transposase complex with DNA
Other DDE conserved transposases?
• Further phylogenetic analysis
References

- Boehm et al. One of three nuclear localization signals of maize Activator (Ac) transposase overlaps the DNA-binding domain. 1995 Plant J


- Kunze et al. Dominant transposition-deficient mutants of maize Activator (Ac) transposase. 1993 PNAS

- Rubin et al, Structure and Evolution of the HAT transposon sumperfamily. 2000 Genetics

- Zhou et al. Transposition of hAT elements links transposable elements and V(D)J recombination. 2004 Nature