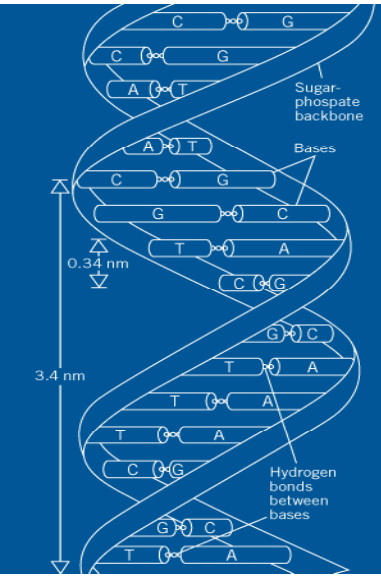
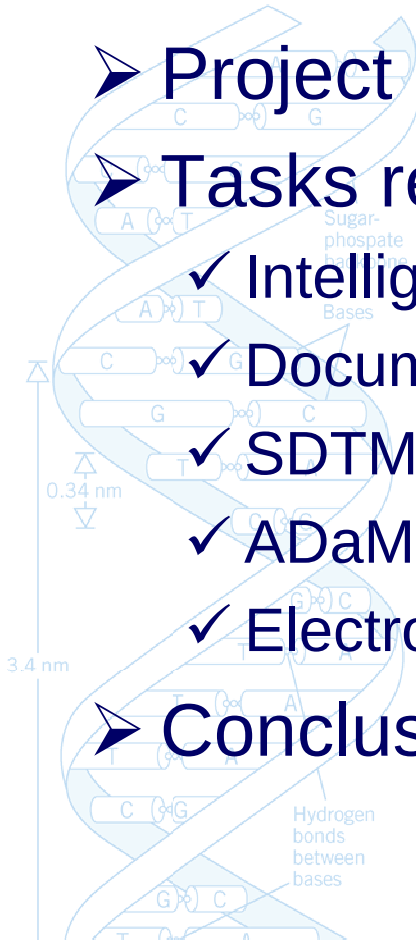


CDISC Implementation on a RA Project Partnership

Patricia Gerend, Olivier Leconte, Chris Price, Michelle Zhang

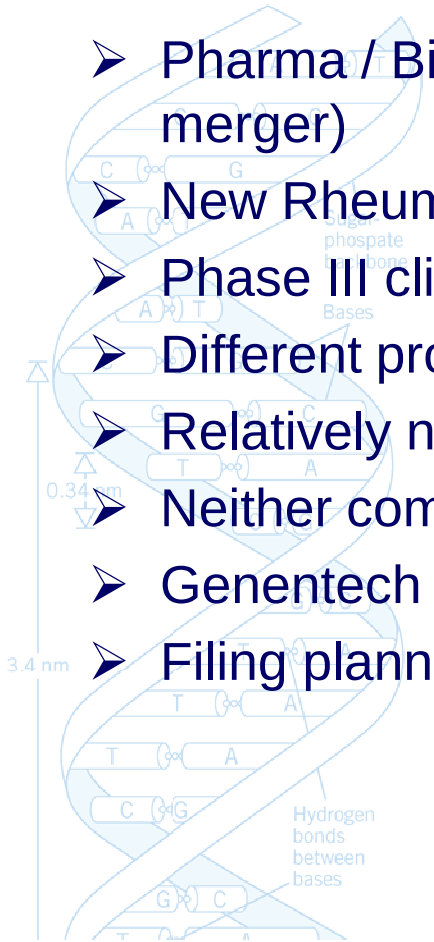
Agenda

- Project Background
- Tasks required for implementation
  - ✓ Intelligence Gathering
  - ✓ Documenting standards
  - ✓ SDTM
  - ✓ ADaM
  - ✓ Electronic Submission
- Conclusions



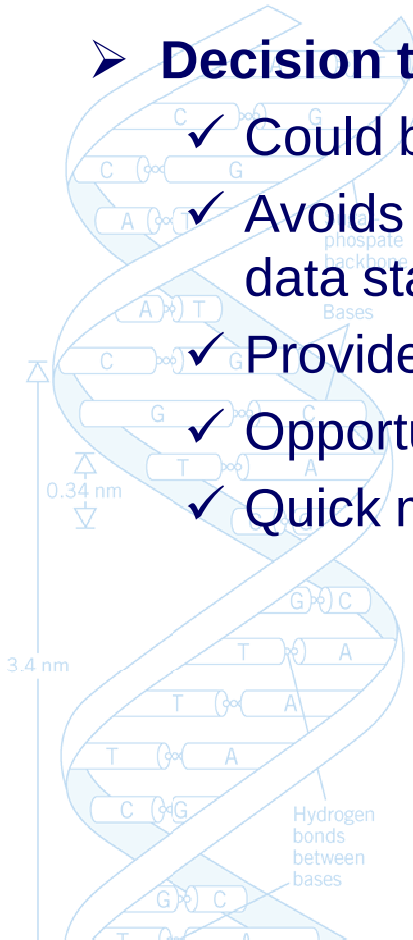
Project Background

- Pharma / Biotech Collaboration: Roche and Genentech (pre-merger)
- New Rheumatoid Arthritis molecule
- Phase III clinical studies getting started
- Different proprietary data standards at each company
- Relatively new industry standard of CDISC
- Neither company had production/filing CDISC experience
- Genentech had performed 2 pilot CDISC projects
- Filing planned for Q3/Q4 2010



Project Background

- **Decision to use CDISC 11/2007**
  - ✓ Could be required by FDA at submission time
  - ✓ Avoids time and hassle of dealing with each other's proprietary data standards
  - ✓ Provides growth opportunity for project team
  - ✓ Opportune timing since project just getting started
  - ✓ Quick management buy-in



Project Background

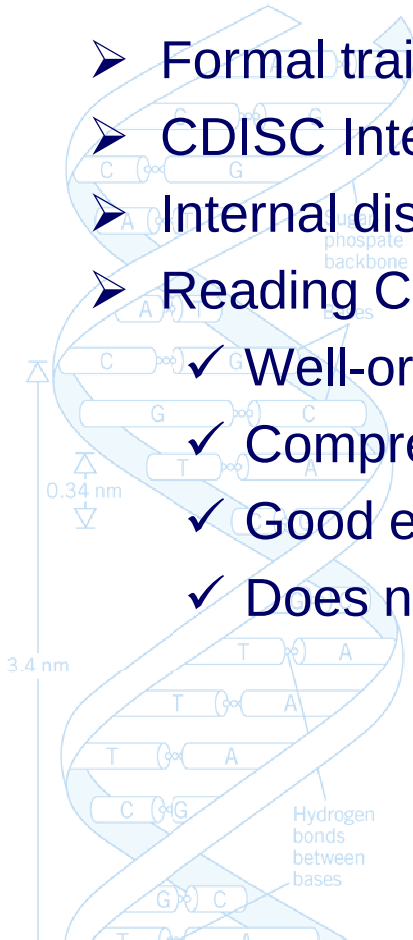
- **Clarify decision making, roles/responsibilities, and accountability**
  - ✓ Genentech to handles data decisions (SDTM and ADaM) since responsible for FDA filing
  - ✓ Roche handles process and systems decisions since work done on Roche systems
- **Communicate, communicate, communicate**

Tasks Required for CDISC Implementation

- Intelligence gathering
- Documenting standards
- SDTM
  - ✓ Modeling of CRFs
  - ✓ Controlled Terminology
  - ✓ Conversion Specifications
  - ✓ Conversions
- ADaM
  - ✓ Analysis Database Design
  - ✓ Metadata and Specifications Structures
  - ✓ Derivations
- Electronic submission to FDA

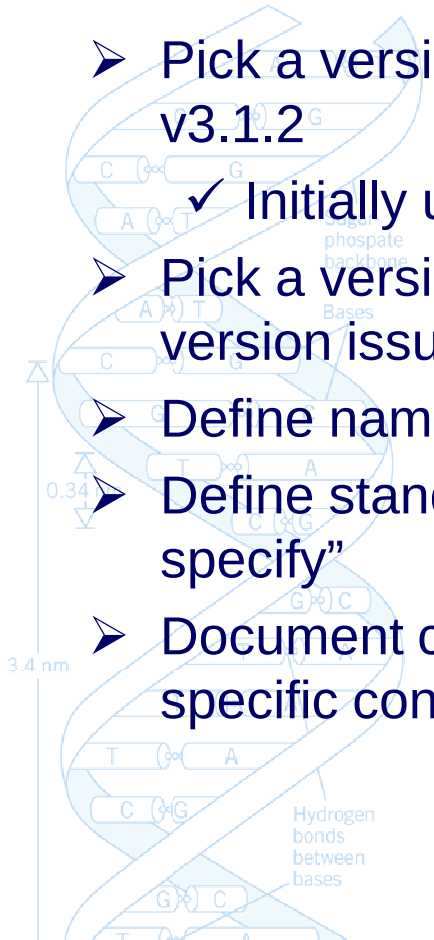
Intelligence Gathering

- Formal training
- CDISC Interchanges and local area groups
- Internal discussions
- Reading CDISC guides
  - ✓ Well-organized
  - ✓ Comprehensive
  - ✓ Good examples
  - ✓ Does not address all modeling situations



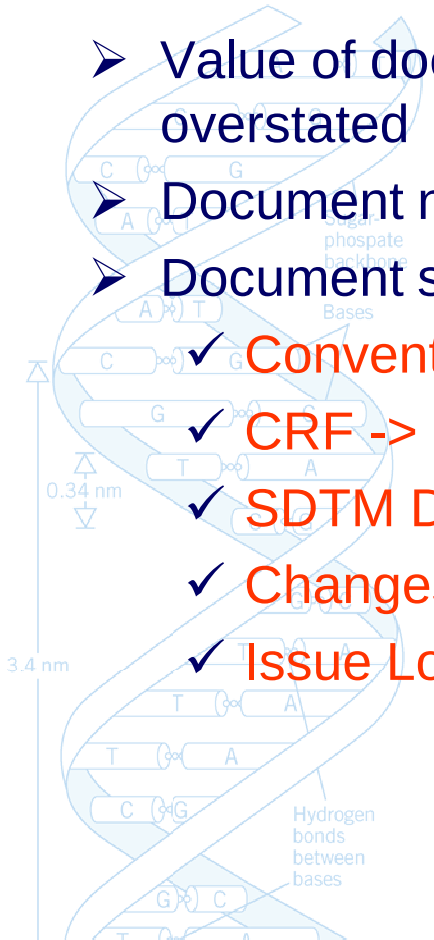
SDTM Modeling

- Pick a version of the CDISC SDTM Implementation Guide (IG): v3.1.2
  - ✓ Initially using the draft version
- Pick a version of the CDISC Controlled Terminology (CT): Recent version issued before first database lock: 7 July 2009
- Define naming conventions for user-defined data domains
- Define standard ways to handle non-standard data, such as “Other, specify”
- Document conventions, modeling decisions, changes, project-specific controlled terminology



SDTM Modeling Documentation

- Value of documentation, though sometimes tedious, cannot be overstated
- Document name: SDTM Modeling Information
- Document sections:
  - ✓ Conventions for SDTM Modeling
  - ✓ CRF -> SDTM Domain Map
  - ✓ SDTM Domain -> CRF Map
  - ✓ Changes to Annotations since First Draft
  - ✓ Issue Log



Conventions for SDTM Modeling

- Dates
  - ✓ Findings domains use xxDTC
  - ✓ Interventions and Events domains use xxSTDTC/xxENDTC.
- User-defined domain naming conventions
  - ✓ Xx for Interventions (e.g., XP for previous procedures)
  - ✓ Yx for Events (e.g., YI for Previous Immunizations)
  - ✓ Zx for Findings (e.g., ZJ for Tender/Swollen Joint Counts)
- A Controlled Terminology spreadsheet for the project was maintained
- Handling of “Other, specify” situations
  - ✓ As part of a pre-specified list of interventions/events → use xxPRESP and xxOCCUR
  - ✓ Multiple related responses → use Findings About (FA)

Controlled Terminology

- Two Controlled Terminology (CT) documents:
  - ✓ CDISC organization
  - ✓ Project
- Identify which version from CDISC organization to use across project
- Identify and document terms specific to project to maintain consistency across studies
- Map project values to CDISC CT where they exist
- Put original values into xxORRES or SUPPQUAL if they differ substantially from CT values
- Remember to check if a CDISC CT value list is extensible
- Identify a Clinical Scientist to use for input into mappings from original to CT values

SDTM Conversion Specifications

- While many conversions are not difficult (e.g., variable re-names), some are, so documentation is helpful
- Set up spreadsheet containing list of all possible variables in the domain and algorithms for populating them

STUDYID	PEPE.STUDY
DOMAIN	“PE”
USUBJID	Concatenate PEPE.STUDY, PEPE.CRTN, and PEPE.PT separated by hyphens
PESEQ	Unique sequence number of PE observation per subject. Create on each record sequentially.

SDTM Conversion SAS Programs

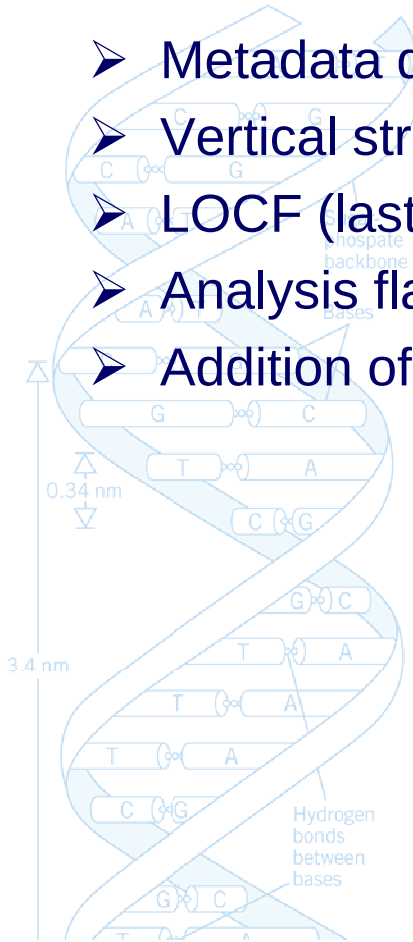
- Base SAS was used to perform the conversions from Oracle Clinical extract data to SDTM
- Advantages over GUI tool used by non-team members
  - ✓ Project team can see entire picture of data derivations
  - ✓ Project team can participate in conversions
  - ✓ All data conversions/derivations are in one programming language with programs residing in one location to facilitate audit trail

ADaM General Decisions

- Used draft versions ADaM Modelv2.1 and ADaM Implementation Guide v1.0
- Decided to go with ADaM basic data structure for efficacy data to experience the CDISC process in its entirety
- Decided to go with proprietary standards with CDISC names for safety to facilitate standard safety reporting

ADaM Challenges

- Metadata documentation
- Vertical structures
- LOCF (last observation carried forward) derivations
- Analysis flags
- Addition of rows versus columns

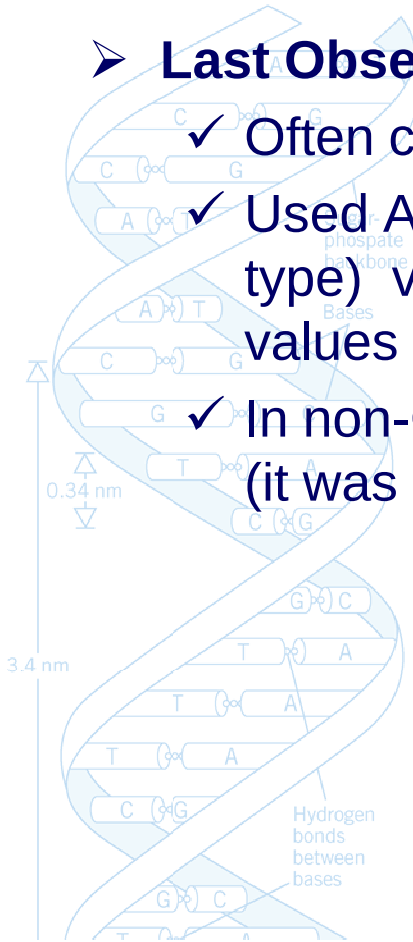


ADaM Specifications

- The following 2-table format was used:
  - ✓ 1-Data List document
  - ✓ 2-Variable List document
  - ✓ Value-level derivation info was embedded into the variable derivation cells
  - ✓ Familiar to FDA reviewers
- Consideration of a 3-table format for future:
  - ✓ 1-Data list document
  - ✓ 2-Variable list document
  - ✓ 3-Value list document
  - ✓ FDA will become more familiar with this in time

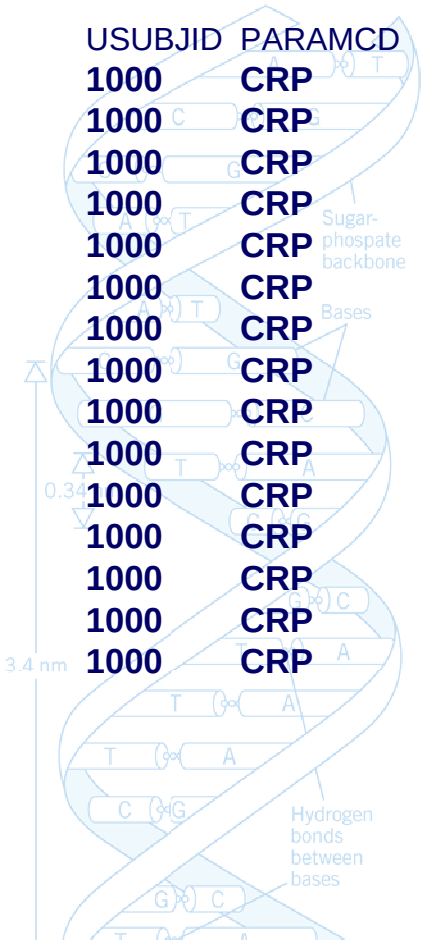
ADaM Derivations Example

- **Last Observation Carried Forward (LOCF)**
  - ✓ Often complicated regardless of data structure
  - ✓ Used ADaM AVAL (analysis variable) and DTYPE (derivation type) variables together to identify observed and LOCF'ed values
  - ✓ In non-CDISC horizontal structures, only 1 variable was needed (it was called LOCF)



ADaM Example

USUBJID	PARAMCD	AVISIT	AVAL	BASE	CHG	DTYPE	ANL1FL(LOCF)	ANL2FL(obs)
1000	CRP	SCREEN	49.9	76.1	-26.2			
1000	CRP	BASELN	76.1	76.1	0		Y	Y
1000	CRP	WK 2	40.6	76.1	-35.5		Y	Y
1000	CRP	WK 4	23.9	76.1	-52.2		Y	Y
1000	CRP	WK 8	22.6	76.1	-53.5			
1000	CRP	WK 8	18.7	76.1	-57.4		Y	Y
1000	CRP	WK 12		76.1				
1000	CRP	WK 12	18.7	76.1	-57.4	LOCF	Y	
1000	CRP	WK 16	14.7	76.1	-61.4		Y	Y
1000	CRP	WK 24	12.8	76.1	-63.3		Y	Y
1000	CRP	WK 32	10.1	76.1	-66.0		Y	Y
1000	CRP	WK 40		76.1				
1000	CRP	WK 40	10.1	76.1	-66.0	LOCF	Y	
1000	CRP	WK 48		76.1				
1000	CRP	WK 48	10.1	76.1	-66.0	LOCF	Y	



ADaM Analysis Flags

- Many different ways of implementing ADaM model
- Had to decide between creating analysis flags for all reasonable analyses or for just those pre-specified: → created all that seemed reasonable
- Example analysis flag: ANL1FL indicates LOCF, excluding rescue and withdrawal
- Decided to have ANLxFL represent same concept across all ADaM datasets, even though this means the value of x is not necessarily sequential in each dataset
  - ✓ Example: ADDS1 contains ANL1FL, ANL2FL, ANL3FL, ANL4FL; ADDS2 contains ANL1FL and ANL4FL
- ADaM model may still be evolving to handle more cases

ADaM Addition of Rows Versus Columns

- Add a new column for a parameter-invariant functions of AVAL (analysis value) or BASE (baseline value) on the same row
  - ✓ "Parameter-invariant" means the function does not change from parameter to parameter and the meaning of the function is the same on all rows
  - ✓ Example: Change from Baseline
- Add a new row for functions that involve more than one parameter or that require a new parameter
  - ✓ Example: Total number of tender joints is derived from each individual joint score, so total number is a new parameter and a new row
  - ✓ Example: LOCF imputation of missing values is put into a new row

Electronic Submission to FDA

- **SDTM**
  - ✓ Followed SDTM-IGv3.1.2 to the best of our abilities
  - ✓ Must still evaluate SDTM structure
  - ✓ Use Phase Forward's WebSDM product to evaluate SDTM structure adherence
  - ✓ define.xml will be generated for filing purposes
  - ✓ Will also generate define.pdf to accommodate reviewers
  - ✓ Will submit dataset list, variable list, and controlled terminology
- **ADaM**
  - ✓ define.pdf, but not define.xml will be generated and submitted
  - ✓ Will submit dataset list and variable list
  - ✓ ADaM structure less stable than SDTM and could change later

Conclusion

- SDTM
  - ✓ First study took 8 months elapsed time
  - ✓ Second study took 3 months elapsed time
  - ✓ Third study took < 1 month elapsed time
- ADaM
  - ✓ First study took 4 months elapsed time
  - ✓ Second study took 2 months elapsed time
  - ✓ Third study took 1 month elapsed time
- CDISC Overall
  - ✓ No haggling over each company's proprietary data structures, so 6 months were saved here

Conclusion

- Results of decision to use CDISC with 2 companies not familiar with its structures
  - ✓ Successful SDTM conversion of 4 studies
  - ✓ Successful ADaM derivation on 3 studies, so far
  - ✓ Intense CDISC learning across both companies
  - ✓ Information to move forward with organization-wide CDISC strategies
- Successes yet to come
  - ✓ Electronic submission deliverables compilation
  - ✓ FDA evaluation of our efforts

Questions?

A
T
C
G
C
G
A
T
Sugar-
phospate
backbone
A
T
Bases
C
G
G
C
T
A
0.34 nm
C
G
G
C
T
A
T
A
T
A
C
G
Hydrogen
bonds
between
bases
G
C
3.4 nm

Acknowledgements

- Genentech
  - ✓ Ian Fleming
  - ✓ Lauren Haworth
  - ✓ Sandra Minjoe (formerly)
  - ✓ Rajkumar Sharma
  - ✓ Peggy Wooster
  - ✓ Susan Zhao
- Roche
  - ✓ Frederik Malfait
- I3Statprobe
  - ✓ Chakrapani Kolluru
- PharmaStat
  - ✓ John Brega
  - ✓ Jane Diefenbach

Contacts

Patricia L. Gerend
Senior Manager, Statistical Programming & Analysis
Genentech, Inc.
South San Francisco, California, USA
gerend@gene.com
650-225-6005

Chris Price
Senior Technical Programmer
Roche Products Limited
Welwyn Garden City, UK
chris.price.cp1@roche.com
+ 44 (0)1707 36 5801

Olivier Leconte
Programming Team Leader
Roche Products Limited
Welwyn Garden City, UK
olivier.leconte@roche.com
+44 (0) 1707 36 5710

Michelle Zhang
Senior Statistical Programmer Analyst
Genentech, Inc.
South San Francisco, CA, USA
zhang@gene.com
650-225-7414