Protein Parts in SBOL

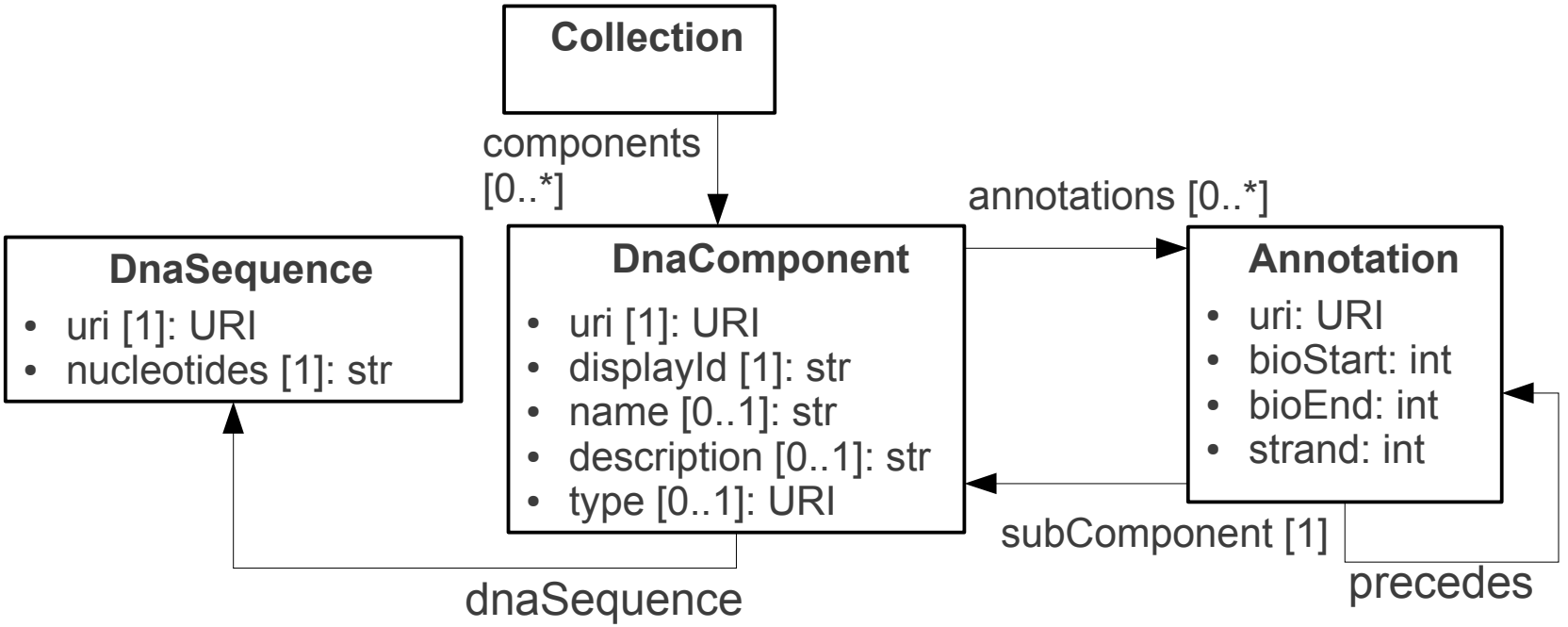
Suggested Extention of the core model

Raik Grünberg
SBOL workshop, San Francisco 05/06/2012

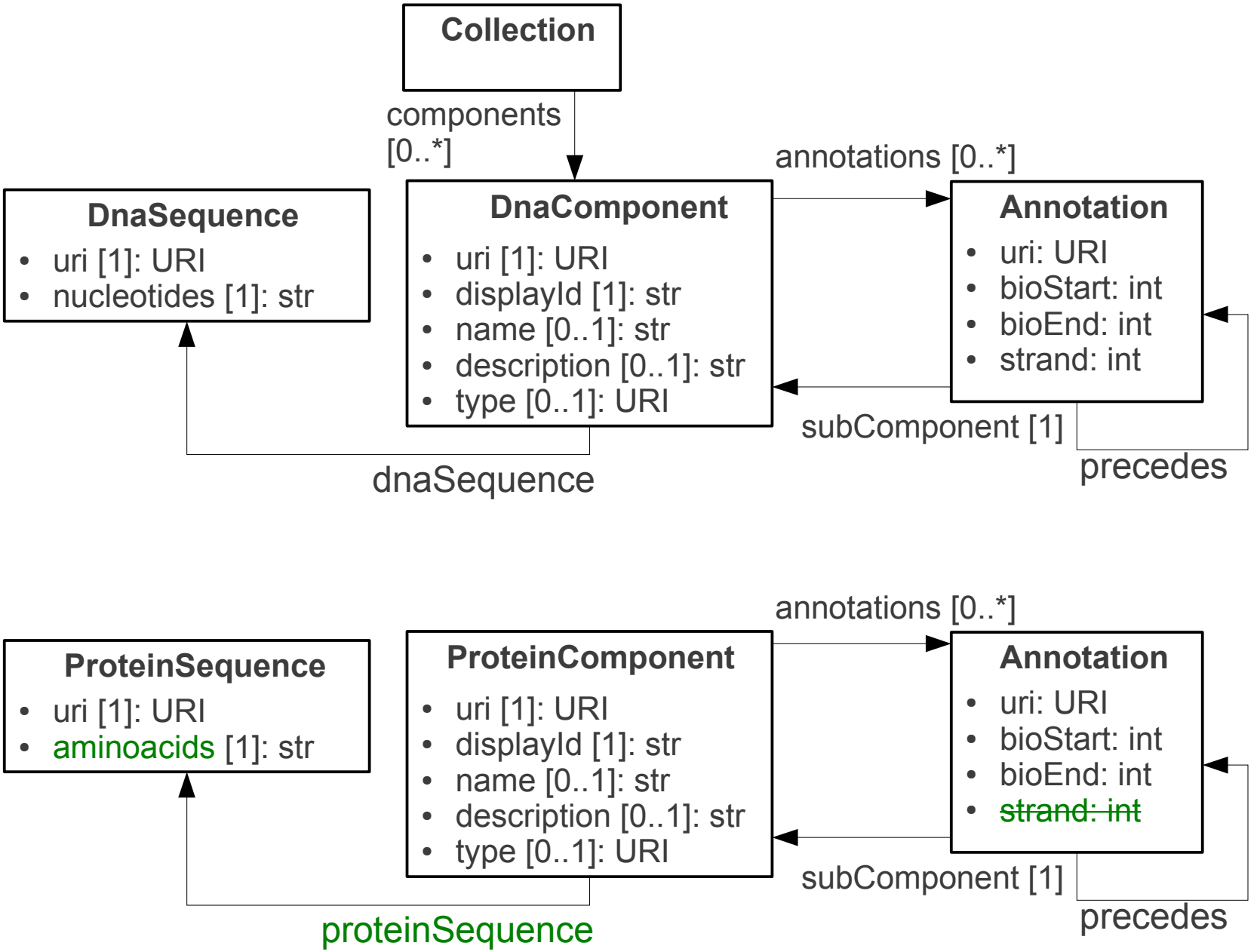
Preview / Summary

```
class DnaComponent(Component):
    """
    Description of a stretch of DNA.
    """
    #: optional sequence
    sequence = models.TextField( help_text='nucleotide sequence',
                                 blank=True, null=True )

    translatesTo = models.ForeignKey( 'ProteinComponent',
                                      blank=True, null=True,
                                      related_name='encodedBy',
        help_text='Protein part this sequence translates to' )

    optimizedFor = models.ForeignKey( 'Chassis',
                                      blank=True, null=True )

    class Meta(Component.Meta):
        pass
```

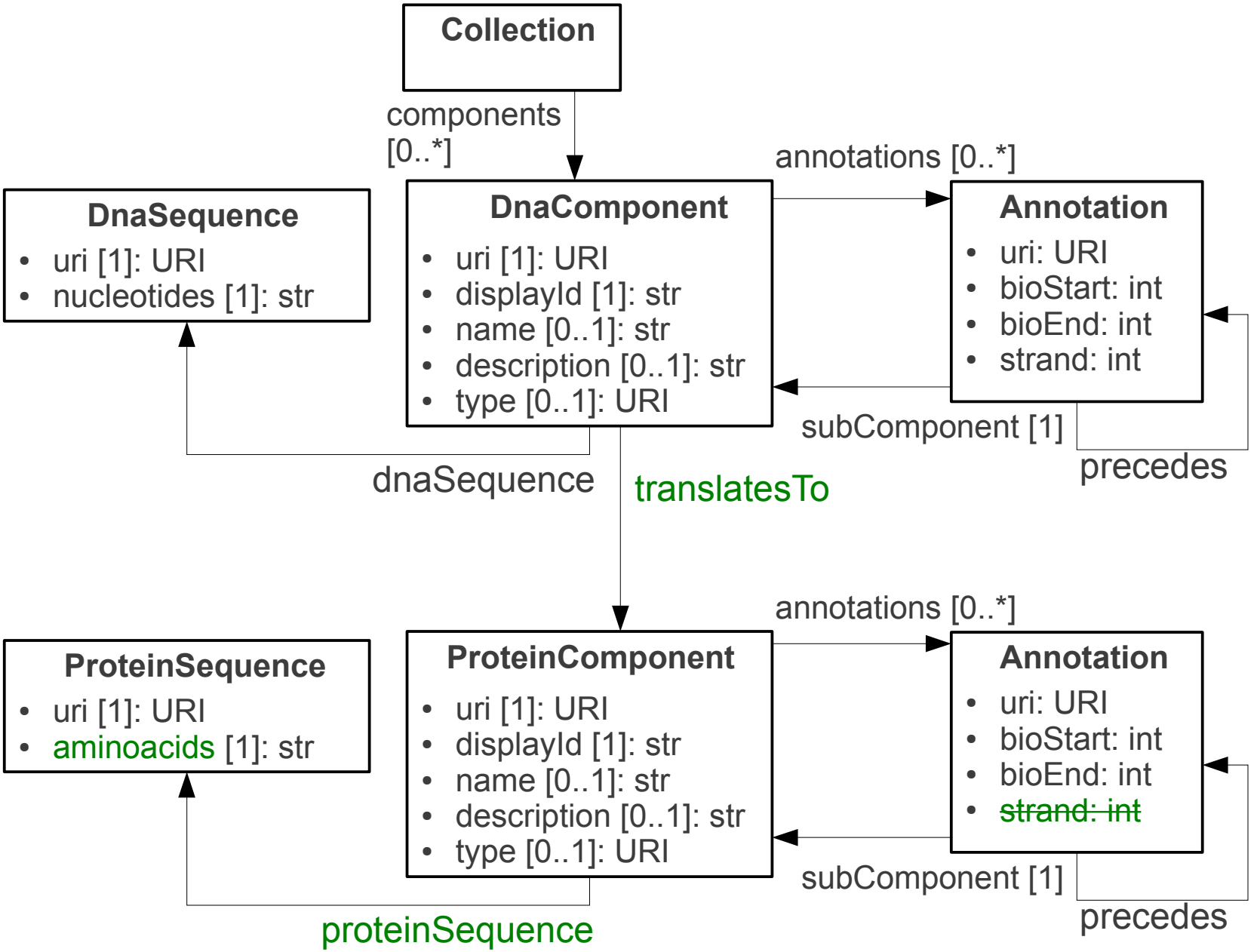
Current core model



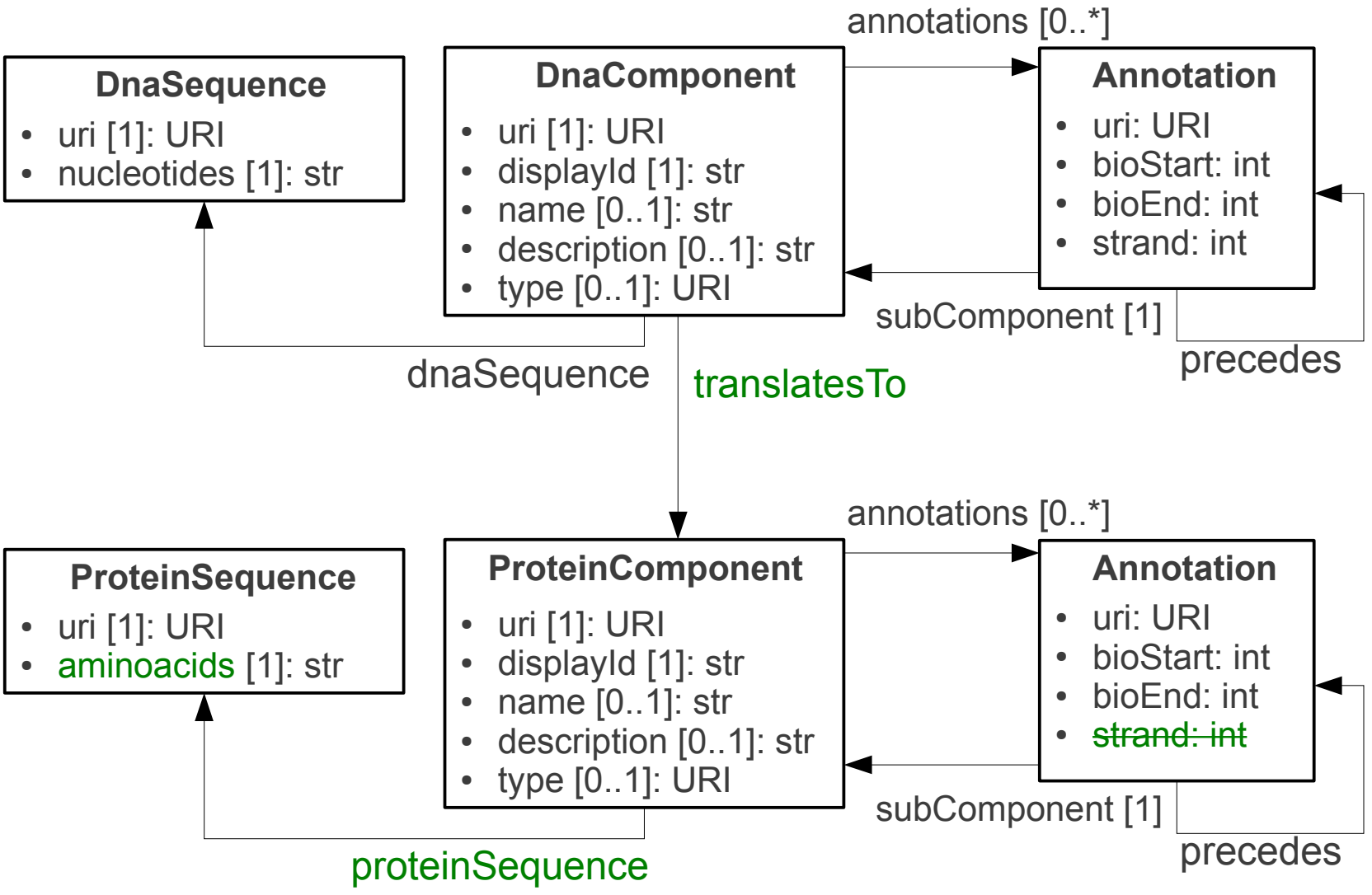
Extended core model



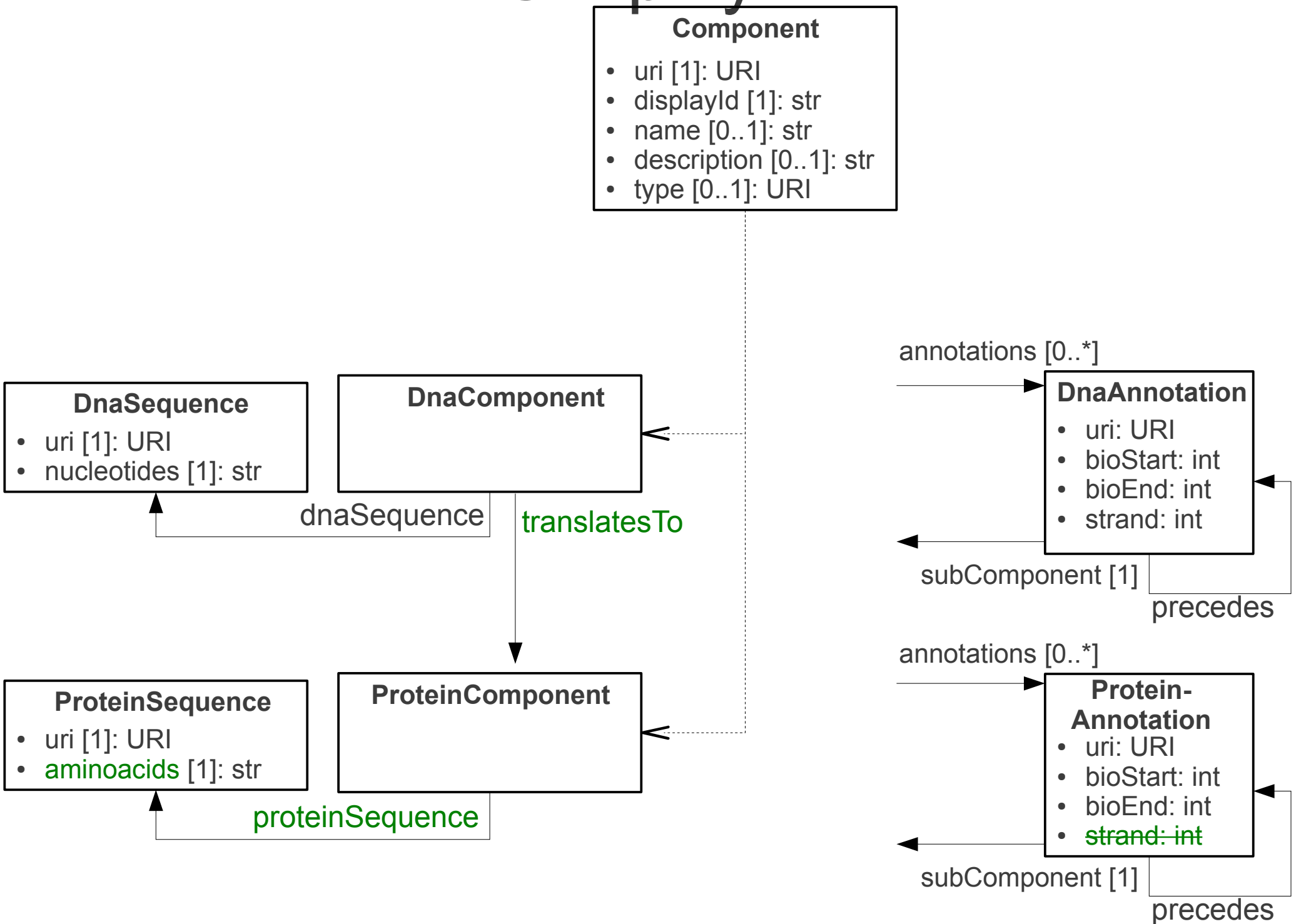
Connect...


Collection
components
[0..*]
DnaSequence
• uri [1]: URI
• nucleotides [1]: str
DnaComponent
• uri [1]: URI
• displayId [1]: str
• name [0..1]: str
• description [0..1]: str
• type [0..1]: URI
annotations [0..*]
Annotation
• uri: URI
• bioStart: int
• bioEnd: int
• strand: int
subComponent [1]
precedes
dnaSequence
translatesTo
ProteinSequence
• uri [1]: URI
• aminoacids [1]: str
ProteinComponent
• uri [1]: URI
• displayId [1]: str
• name [0..1]: str
• description [0..1]: str
• type [0..1]: URI
annotations [0..*]
Annotation
• uri: URI
• bioStart: int
• bioEnd: int
• strand: int
subComponent [1]
precedes
proteinSequence

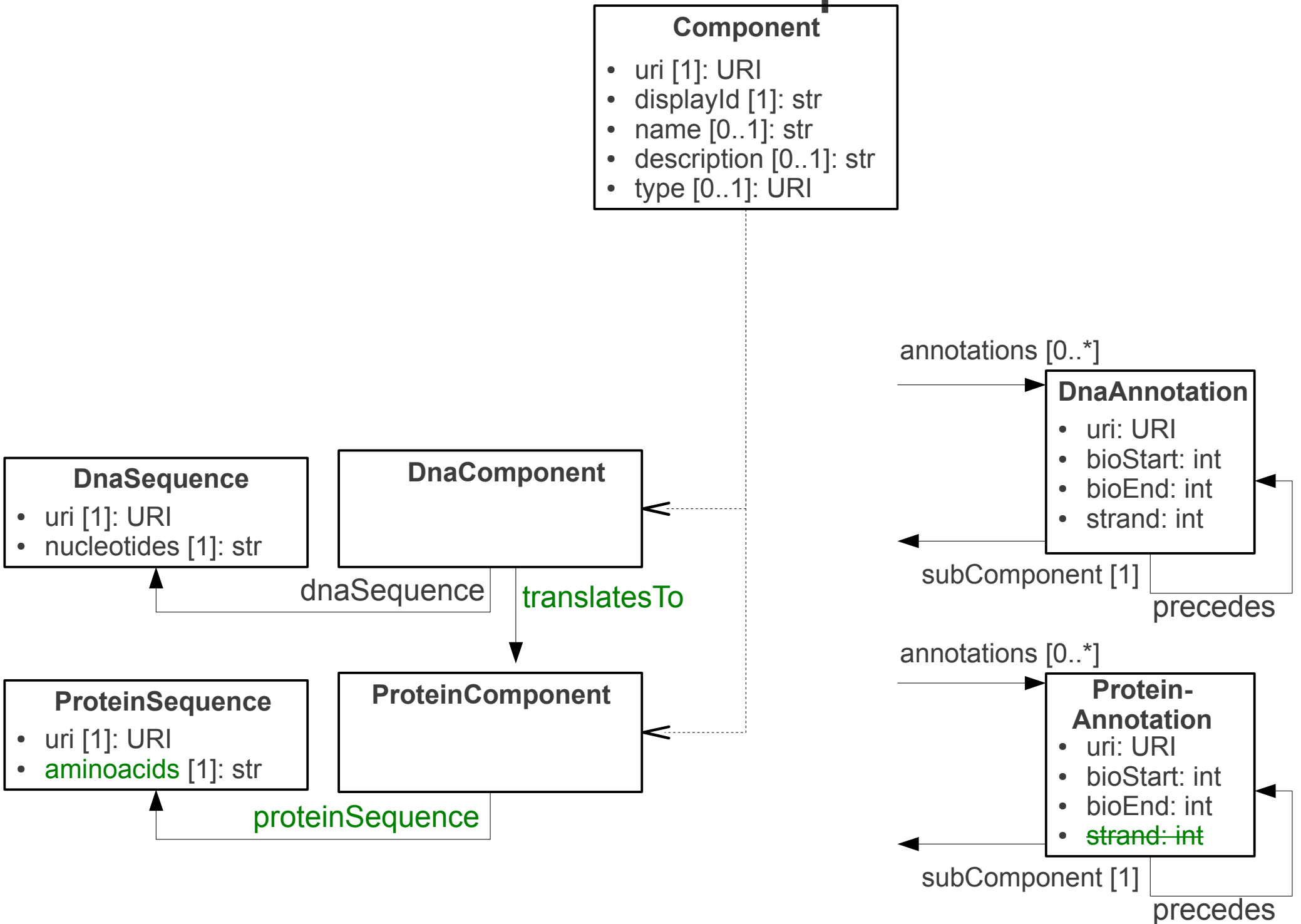
Simplify...


DnaSequence
uri [1]: URI
nucleotides [1]: str
DnaComponent
uri [1]: URI
displayId [1]: str
name [0..1]: str
description [0..1]: str
type [0..1]: URI
annotations [0..*]
Annotation
uri: URI
bioStart: int
bioEnd: int
strand: int
subComponent [1]
precedes
dnaSequence
translatesTo
ProteinSequence
uri [1]: URI
aminoacids [1]: str
ProteinComponent
uri [1]: URI
displayId [1]: str
name [0..1]: str
description [0..1]: str
type [0..1]: URI
annotations [0..*]
Annotation
uri: URI
bioStart: int
bioEnd: int
strand: int
subComponent [1]
precedes
proteinSequence

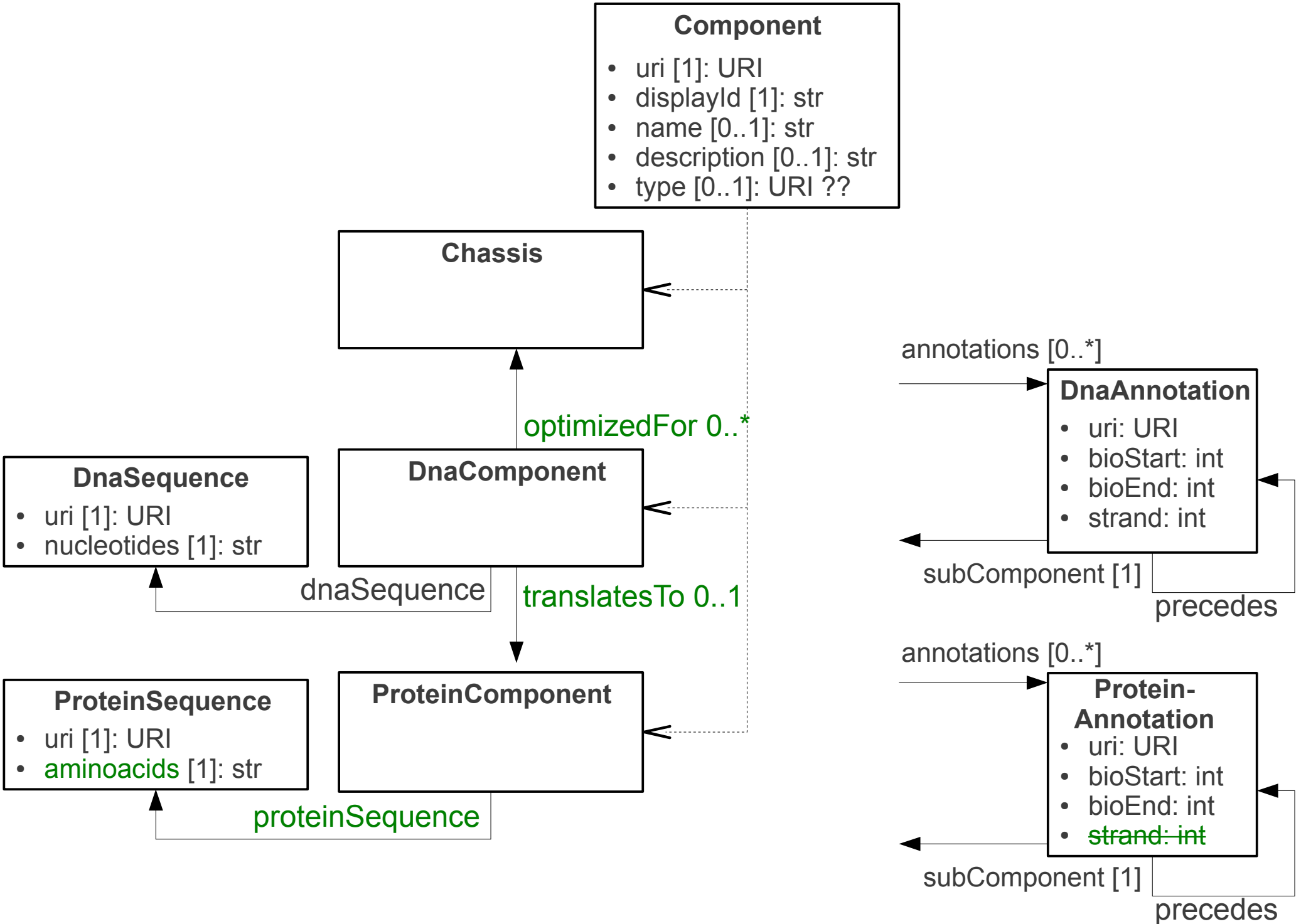
Simplify...


Component
uri [1]: URI
displayId [1]: str
name [0..1]: str
description [0..1]: str
type [0..1]: URI
DnaSequence
uri [1]: URI
nucleotides [1]: str
DnaComponent
dnaSequence
translatesTo
annotations [0..*]
DnaAnnotation
uri: URI
bioStart: int
bioEnd: int
strand: int
subComponent [1]
precedes
ProteinSequence
uri [1]: URI
aminoacids [1]: str
ProteinComponent
proteinSequence
annotations [0..*]
Protein-Annotation
uri: URI
bioStart: int
bioEnd: int
strand: int
subComponent [1]
precedes

Next little step...

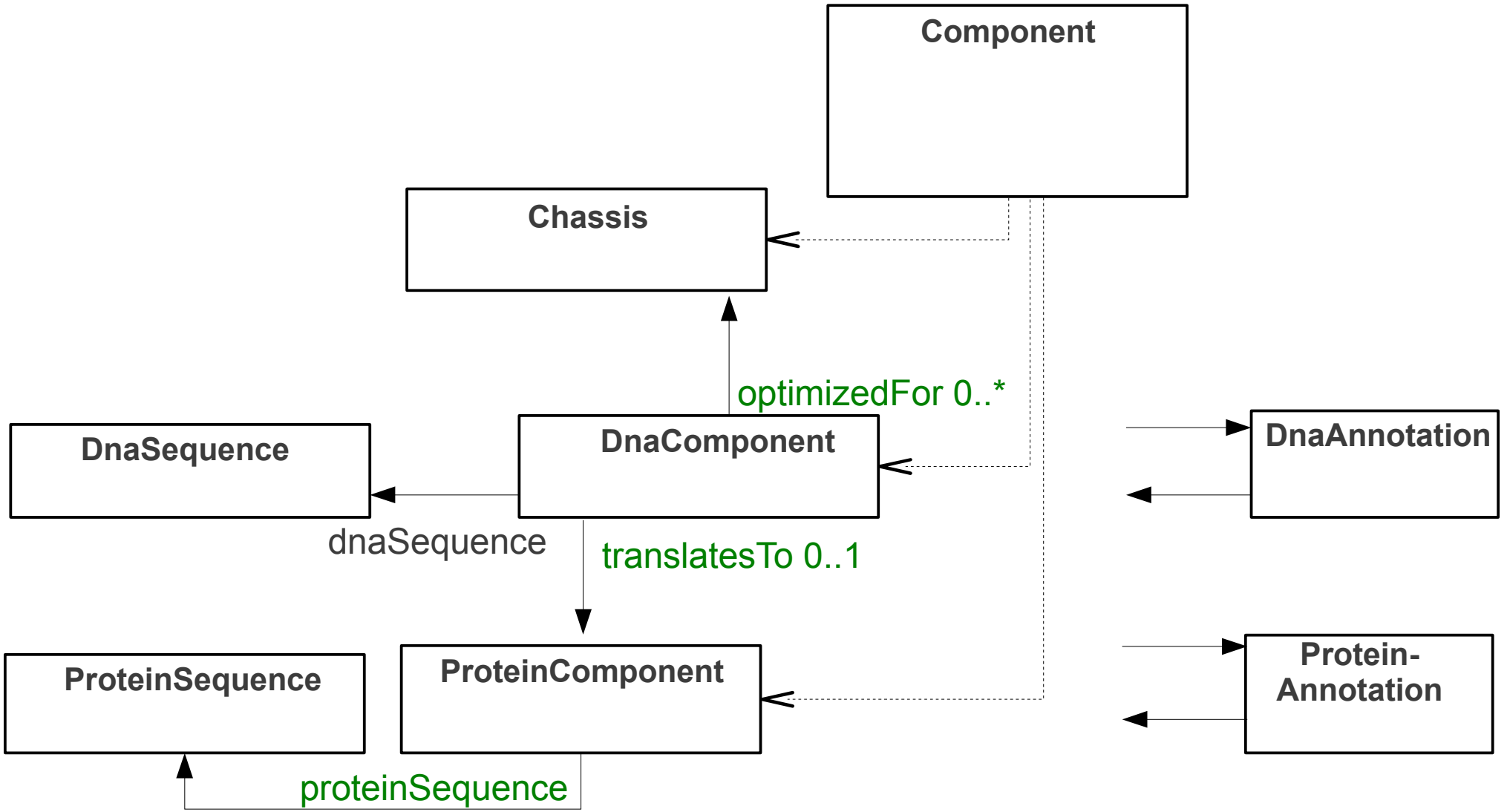


Chassis

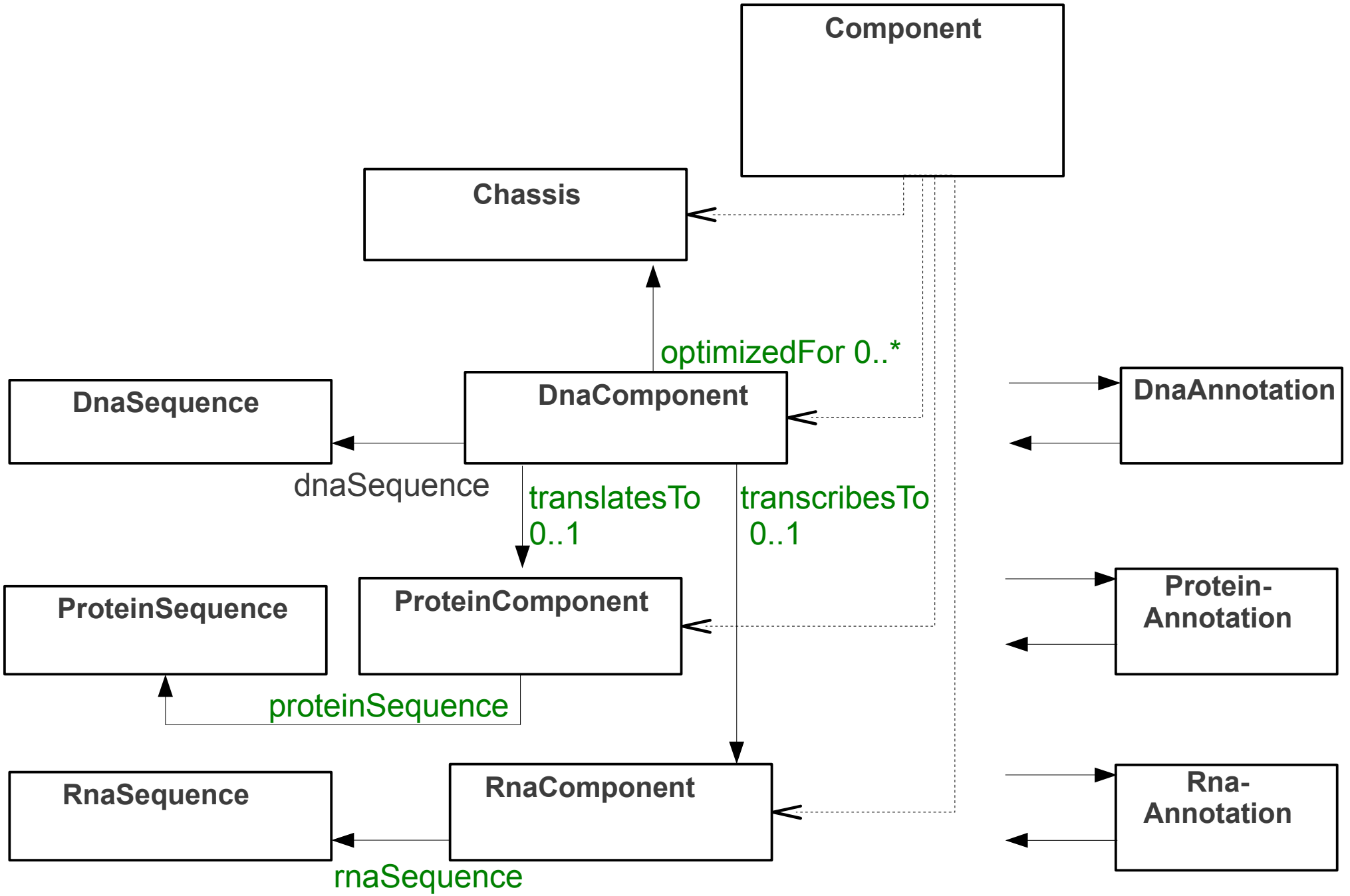

Component
uri [1]: URI
displayId [1]: str
name [0..1]: str
description [0..1]: str
type [0..1]: URI ??
Chassis
optimizedFor 0..*
DnaSequence
uri [1]: URI
nucleotides [1]: str
DnaComponent
dnaSequence
translatesTo 0..1
ProteinSequence
uri [1]: URI
aminoacids [1]: str
ProteinComponent
proteinSequence
annotations [0..*]
DnaAnnotation
uri: URI
bioStart: int
bioEnd: int
strand: int
subComponent [1]
precedes
annotations [0..*]
Protein-Annotation
uri: URI
bioStart: int
bioEnd: int
strand: int
subComponent [1]
precedes

Summary -- outlook

Complete the core...



Complete Core definition

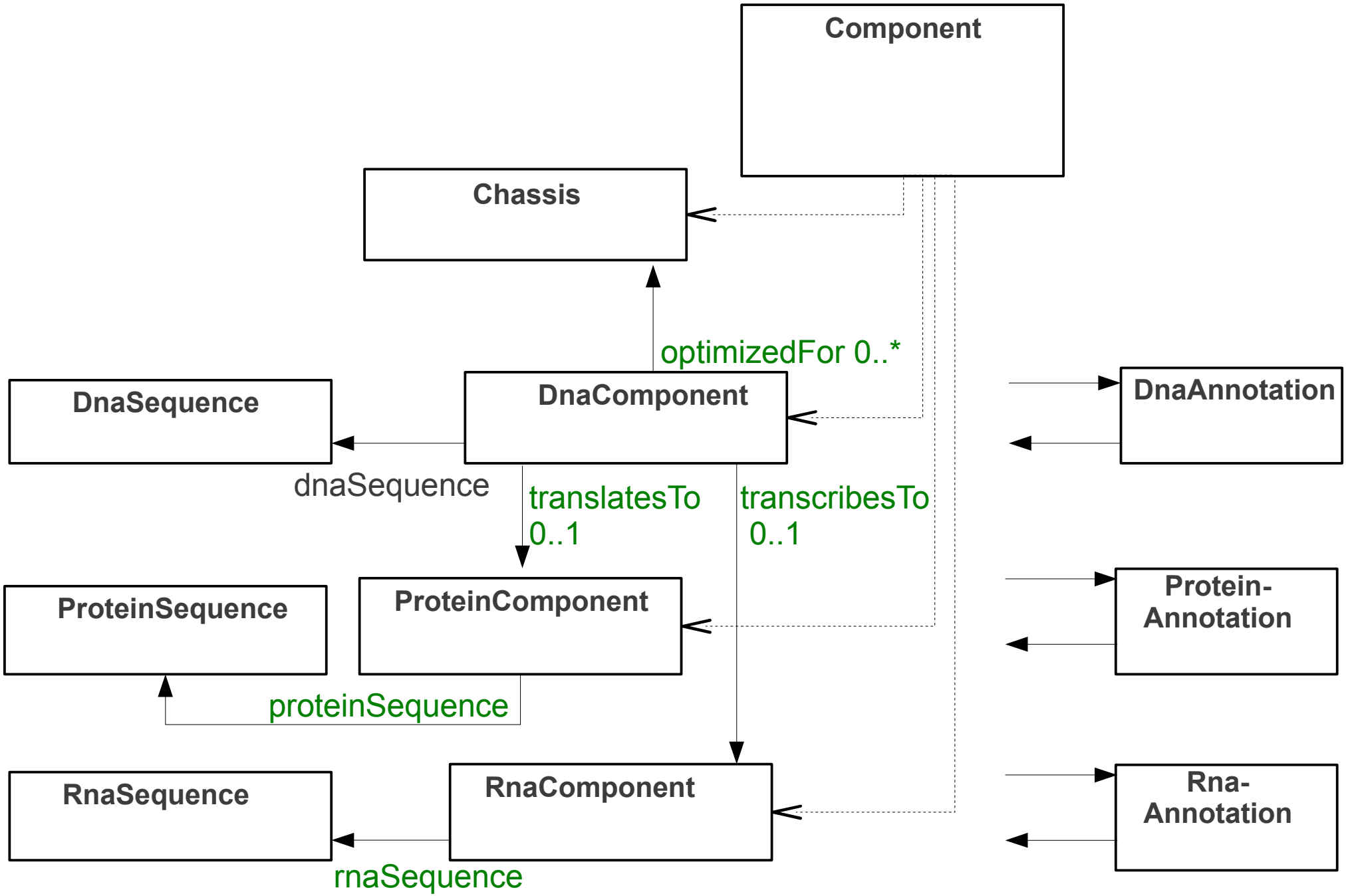


Additional core extension (perhaps not for today :)

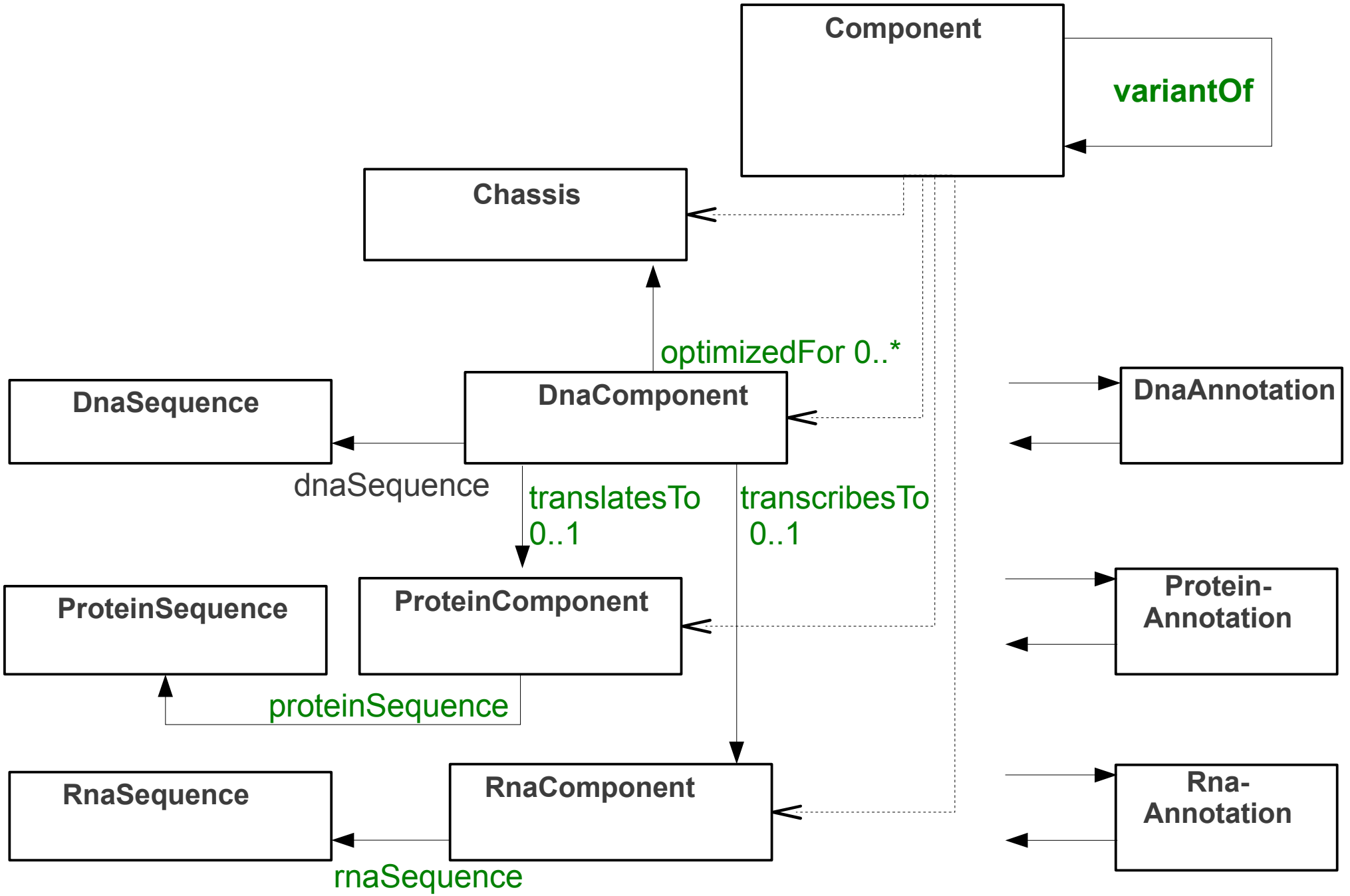
Part inheritence

Classification / Versioning / Optimization variants

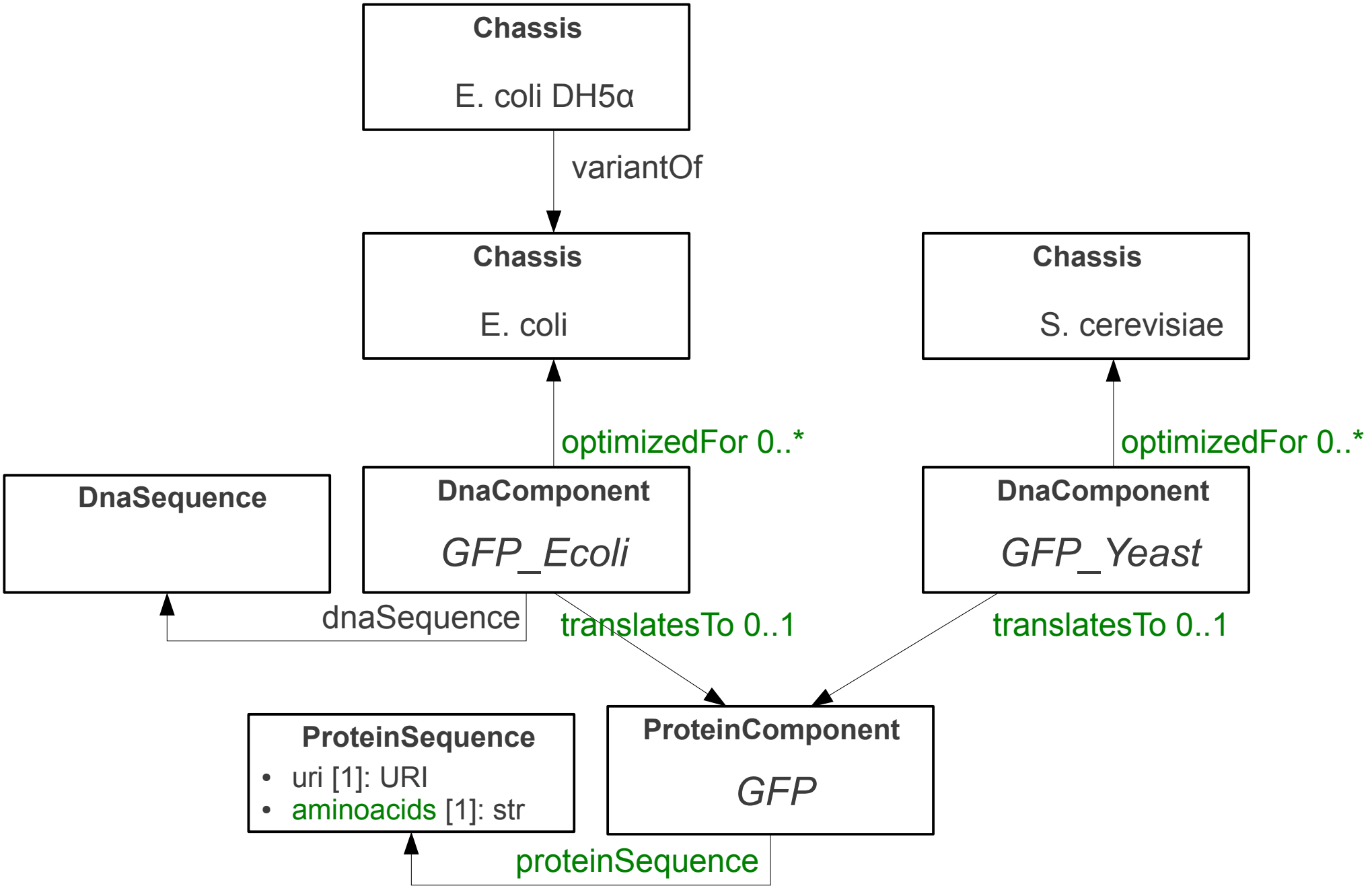
Complete Core definition



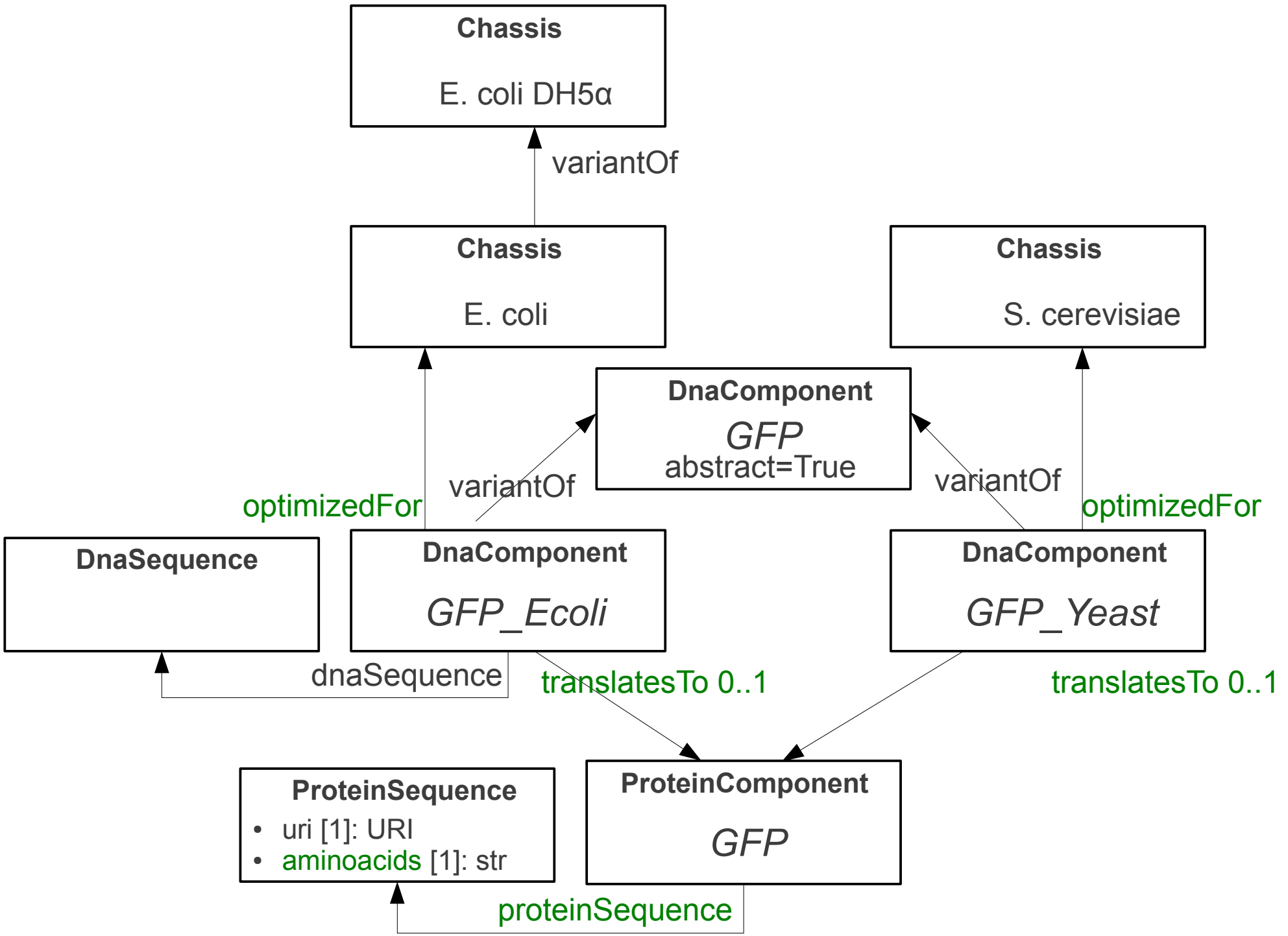
Complete Core definition



Example


Chassis
E. coli DH5α
variantOf
Chassis
E. coli
Chassis
S. cerevisiae
optimizedFor 0..*
optimizedFor 0..*
DnaSequence
DnaComponent
GFP_Ecoli
DnaComponent
GFP_Yeast
dnaSequence
translatesTo 0..1
translatesTo 0..1
ProteinSequence
• uri [1]: URI
• aminoacids [1]: str
ProteinComponent
GFP
proteinSequence

Example


Chassis
E. coli DH5α
variantOf
Chassis
E. coli
Chassis
S. cerevisiae
DnaComponent
GFP
abstract=True
optimizedFor
variantOf
variantOf
optimizedFor
DnaSequence
DnaComponent
GFP_Ecoli
DnaComponent
GFP_Yeast
dnaSequence
translatesTo 0..1
translatesTo 0..1
ProteinSequence
• uri [1]: URI
• aminoacids [1]: str
ProteinComponent
GFP
proteinSequence

Discussion