3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019



3DEXPERIENCE®

BIOVIA Discovery Studio

COMPREHENSIVE MODELING AND SIMULATIONS FOR LIFE SCIENCES

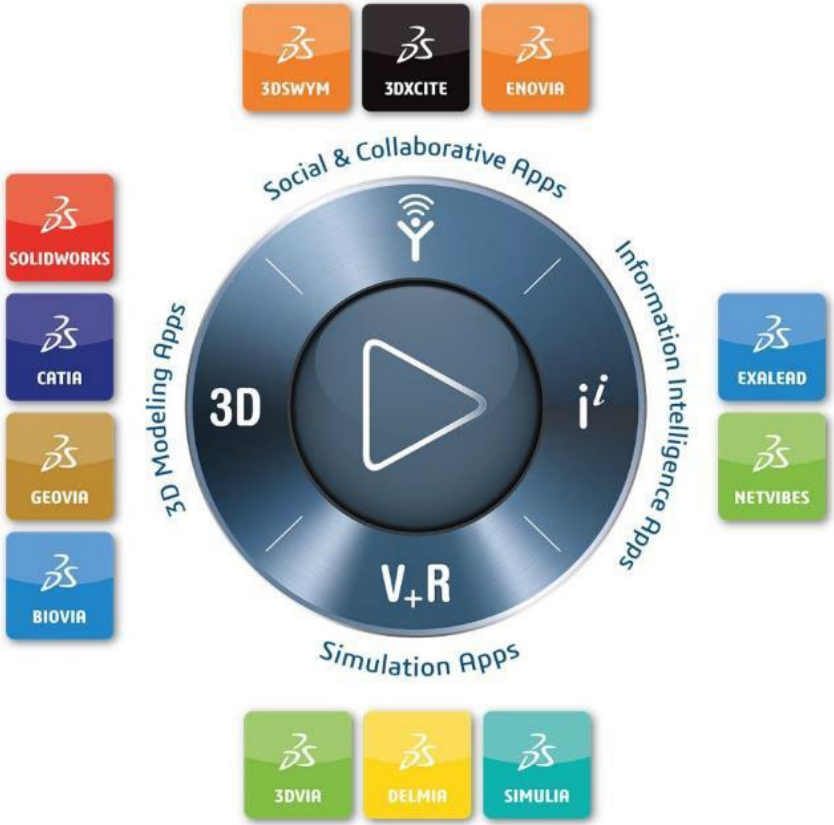
創源生技 分子視算中心

經理 陳冠文 (Gene)

創源生技 GGA

DASSAULT SYSTEMES | The 3DEXPERIENCE® Company

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |



Copyright©2019 GGA Corp., All rights reserved.

Copyright and Disclaimer

- Copyright © 2019 GGA corp. All rights reserved.
  - This presentation and/or any related documents contains statements regarding our plans or expectations for future features, enhancements or functionalities of current or future products (collectively "Enhancements"). Our plans or expectations are subject to change at any time at our discretion. Accordingly, GGA Corp. is making no representation, undertaking no commitment or legal obligation to create, develop or license any product or Enhancements.
  - The presentation, documents or any related statements are not intended to, nor shall, create any legal obligation upon GGA Corp., and shall not be relied upon in purchasing any product. Any such obligation shall only result from a written agreement executed by both parties.
  - In addition, information disclosed in this presentation and related documents, whether oral or written, is confidential or proprietary information of GGA Corp.. It shall be used only for the purpose of furthering our business relationship, and shall not be disclosed to third parties.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

GGA is part of the BIONET Group (訊聯生物科技)



CEO: Christopher Tsai, Ph.D. 蔡政憲 博士

Established: Nov. 2008

Main Product & Service Areas:

1. Genetic Testing
2. Molecular Diagnosis
3. **Scientific Informatics & Bio IT**

IPO Date: September 17, 2012

Stock Ticker: 4160 (Taiwan OTC)

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

About Dassault Systèmes



a **Scientific** company

Combining **Science, Te chnology** and **Art** for a sustainable society



12,400 passionate people

- 112 nationalities
- One global R&D / 53 labs
- Game changing **3D**EXPERIENCE solutions



190,000 enterprise customers

- 12 industries in 140 countries
- >10 million on premise users
- >100 million online users



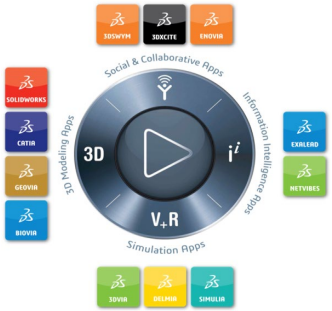
3,500 partners

- Research & Education
- Software & Technology
- Sales & Services



Long-term driven

- Majority shareholder control
- **Revenue**: **$ 2.8 Bn***
- Operating margin: 31.5%*

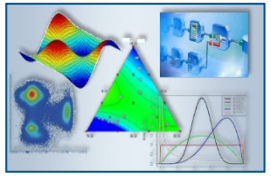


Copyright©2019 GGA Corp., All rights reserved

*Estimated

BIOVIA Perfect Discovery

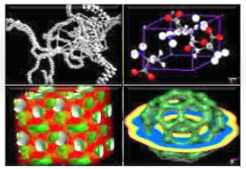
Discover new materials and ingredients in silico to enhance new product offerings



Optimize for Sustainable Performance & Manufacture

Link chemical structure and performance, and optimize trade-offs between multiple requirements

3



Perfect Discovery

Validate Breakthrough Materials



2

Rapidly test and validate virtual material samples that may be too expensive, complicated, or risky to synthesize experimentally

1



Screen Candidate Molecules

Rapidly screen large volumes of molecules to find new leads, and avoid IP conflicts

Copyright©2019 GGA Corp., All rights reserved.

Agenda

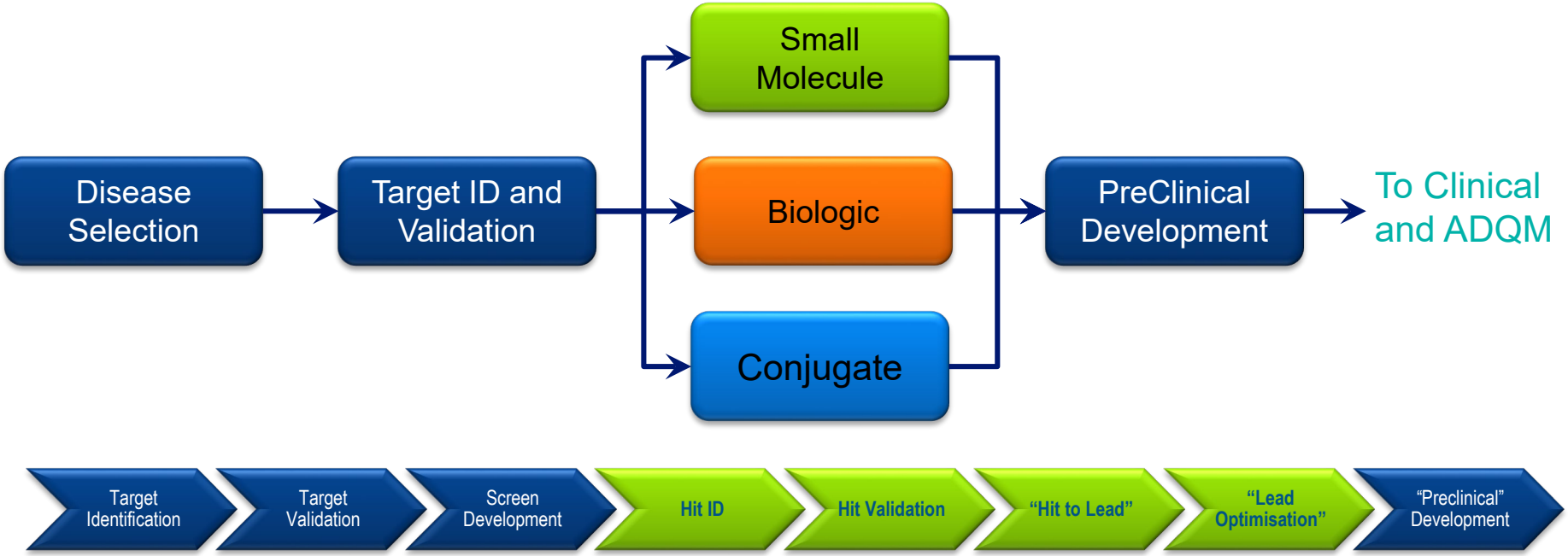
- **Introduction to Predictive Science**
  - Challenges of the Life Science Industry
  - Small Molecules and Macromolecule Solution
  - Hands-on 基本介面操作教學 (實機演練)
- **Discovery Studio: Receptor-Ligand Interactions**

  小分子化合物接合模擬
  - Hands-on操作教學 (實機演練)

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Life Sciences R&D Top Level Workflow



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

“Making” Biotherapeutics Is Different

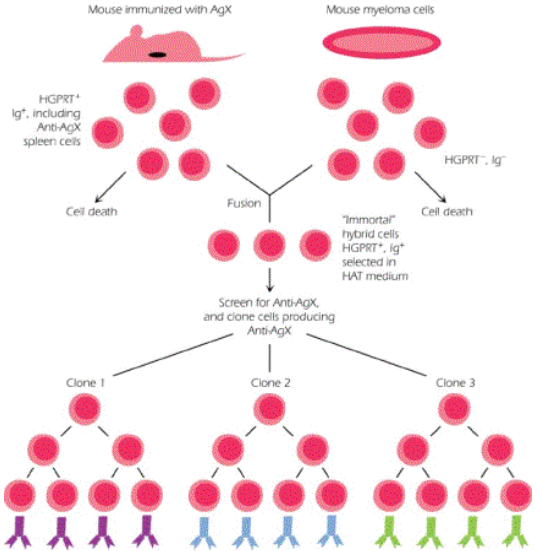
Small Molecule



Chemist synthesizes the molecule

Versus

Biotherapeutics



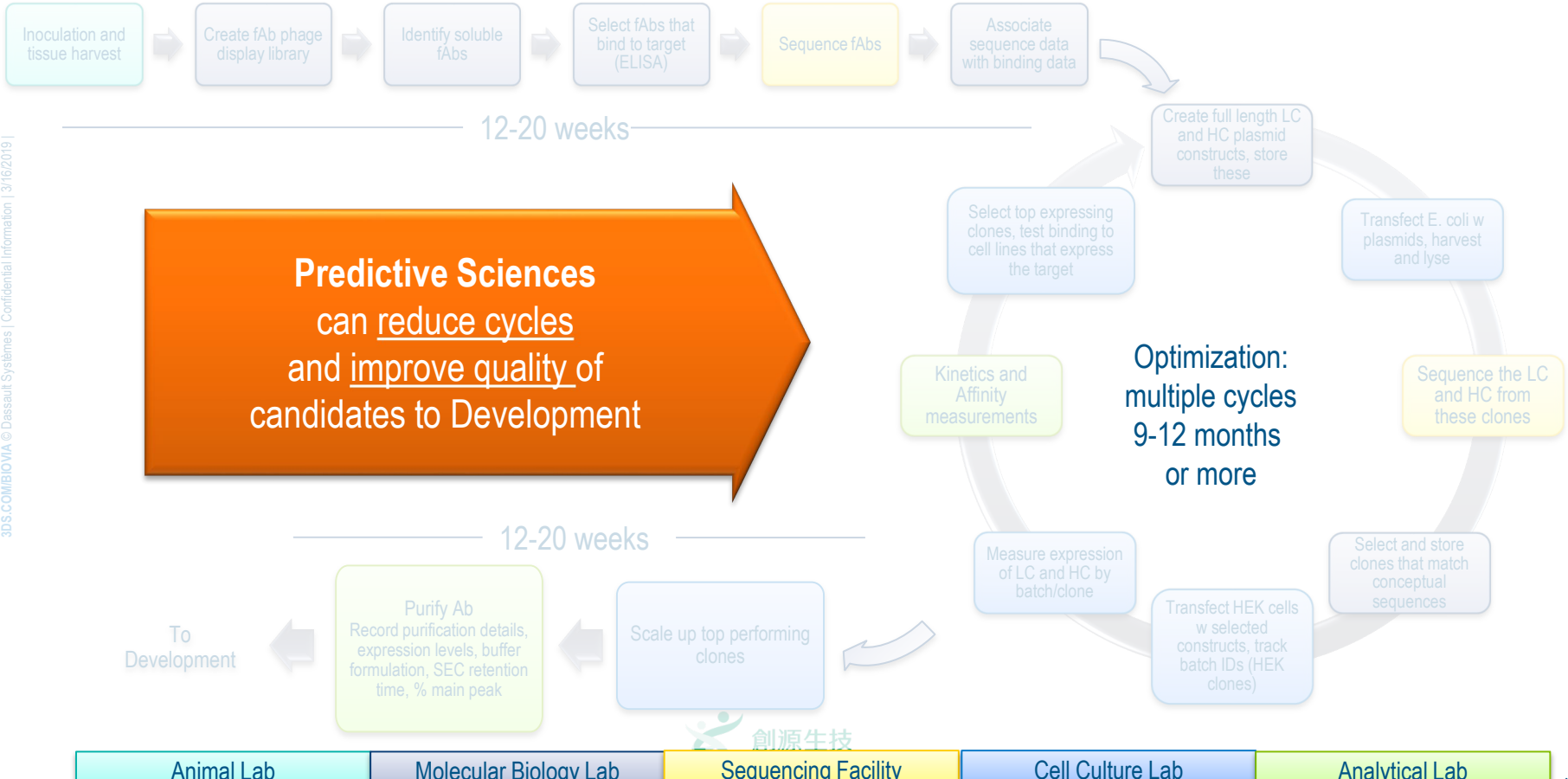
Biological systems make the molecule

Extensive experimentation, random variation and screening intrinsic to the process

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

served.

Biotherapeutic Discovery is Labor and Time Intensive!



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

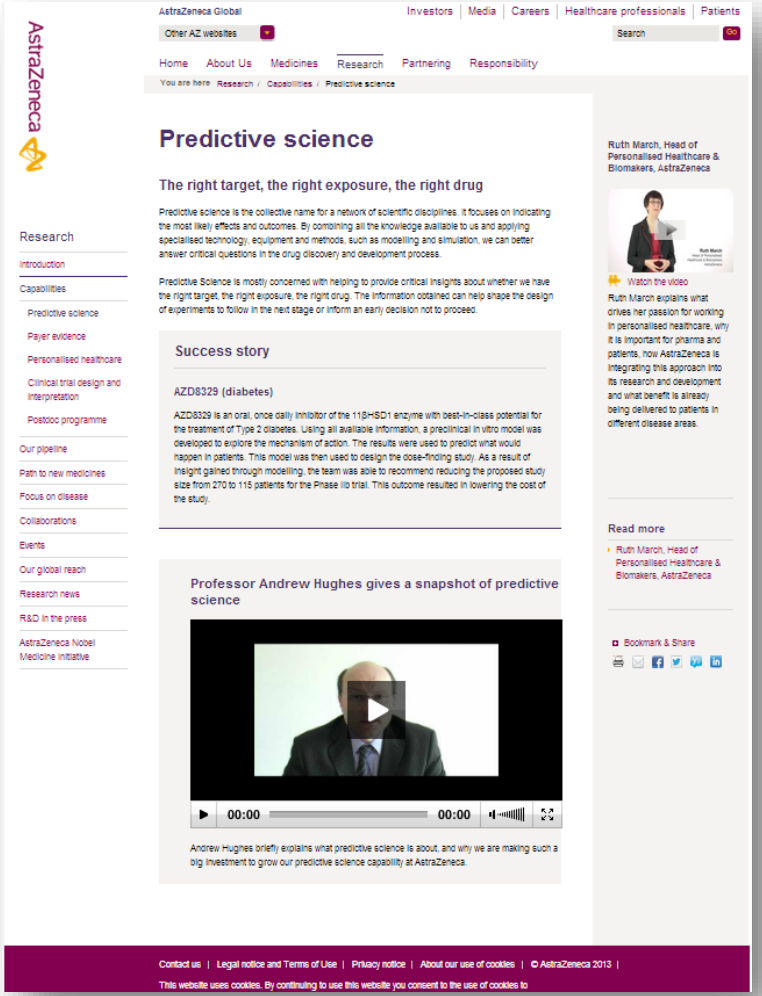
創源生技

Pharma And Predictive Sciences…

- Example: AstraZeneca

> ***"The right target, the right exposure, the right drug"***
>
> *At AstraZeneca, we're spearheading the model-based approach to drug research and development. In our mission to bring life-changing medicines to the global population, we're doubling our investment in Predictive Sciences – so we can tackle illness and disease with even more effective drugs.*

Improve Quality, Save Time, Save Money

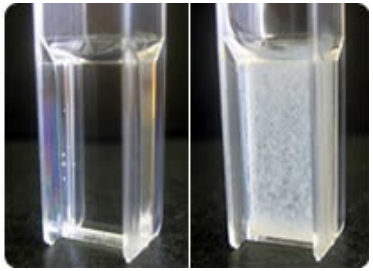


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Biotherapeutics Development Challenges

- Examples of failures in Development include:





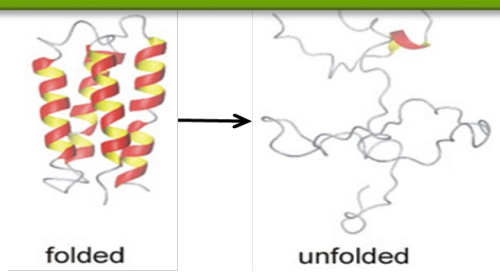


BIOVIA Predictive Sciences Can Help Identify These Problems Earlier
Can Help to Avoid Costly Development Failures



Figure 2. Rash observed with cetuximab therapy before (A) and after (B) treatment with topical steroid. From Sipples[3] with permission. (C) Hyperpigmentation following acneiform eruption with also some new erythematous lesions. From Segaert[4] with permission.

Immunogenicity



Low thermal stability

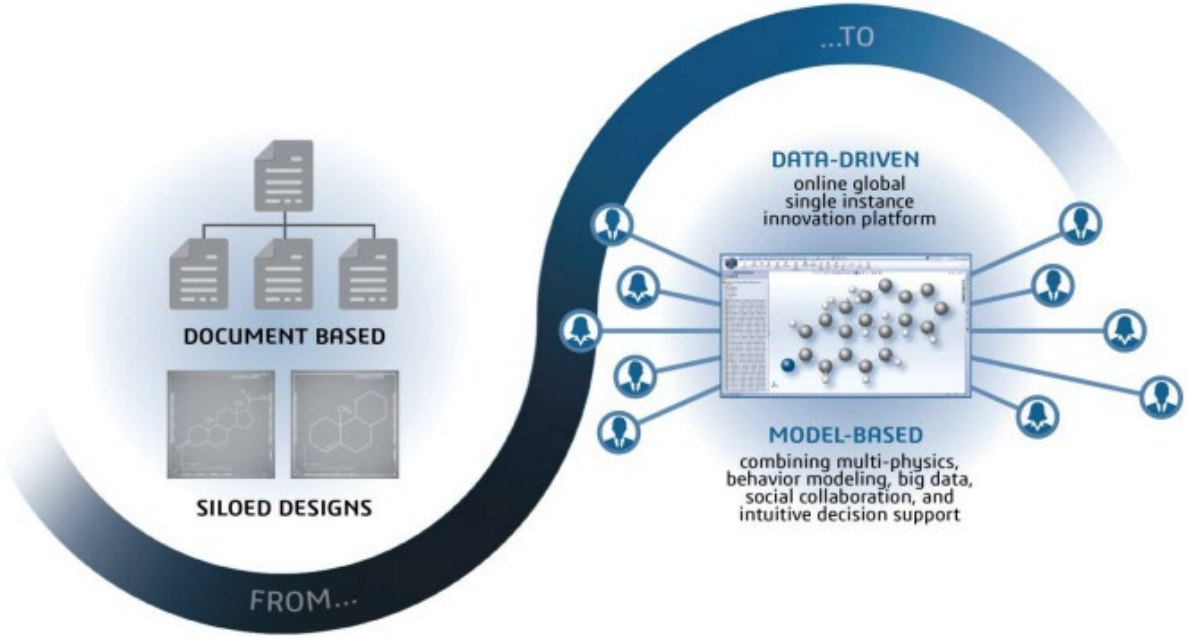


Poor serum half-life

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Generational changes with 3DEXPERIENCE



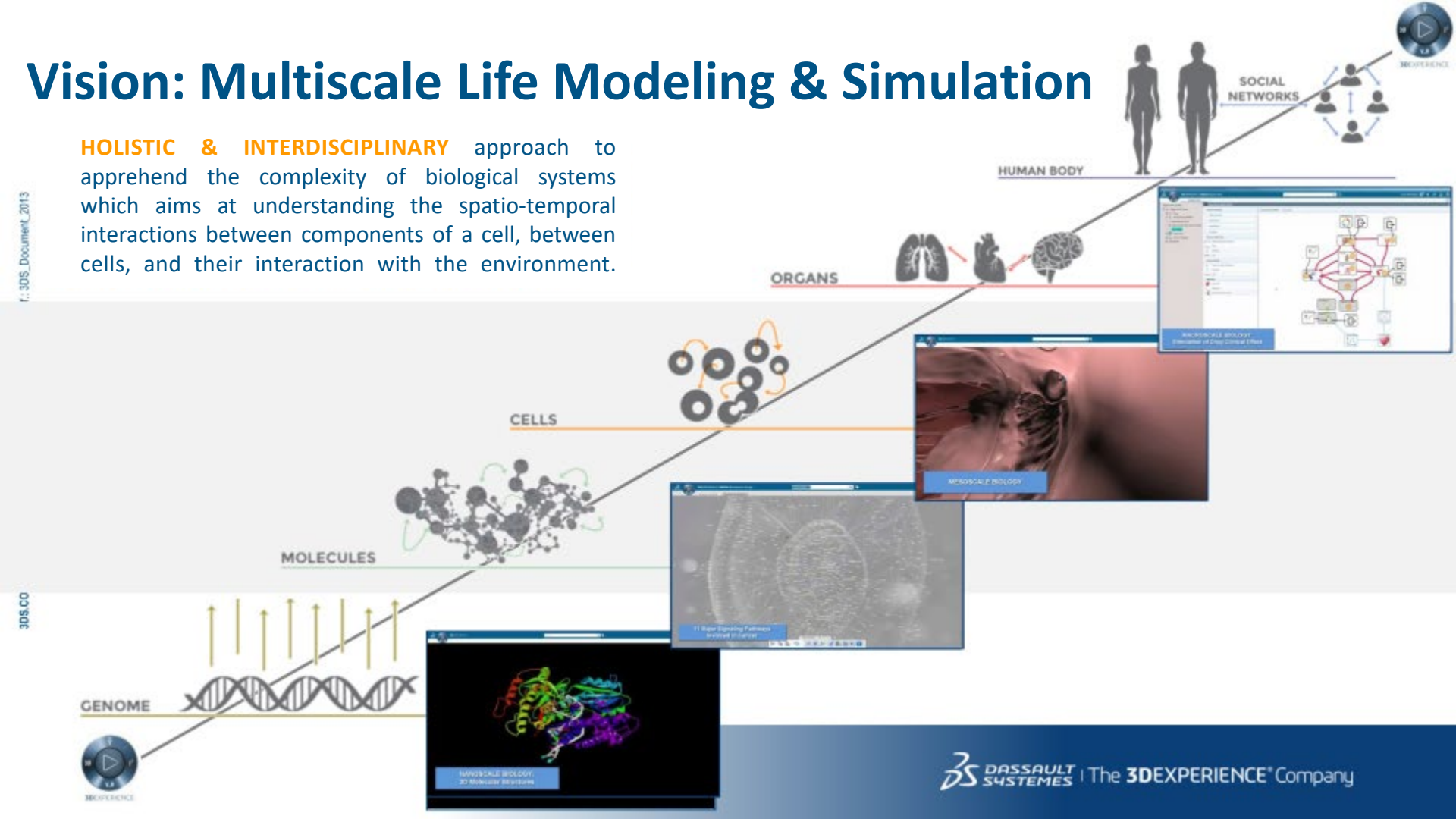
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技
GGA

Copyright©2019 GGA Corp., All rights reserved.

Vision: Multiscale Life Modeling & Simulation

HOLISTIC & INTERDISCIPLINARY approach to apprehend the complexity of biological systems which aims at understanding the spatio-temporal interactions between components of a cell, between cells, and their interaction with the environment.



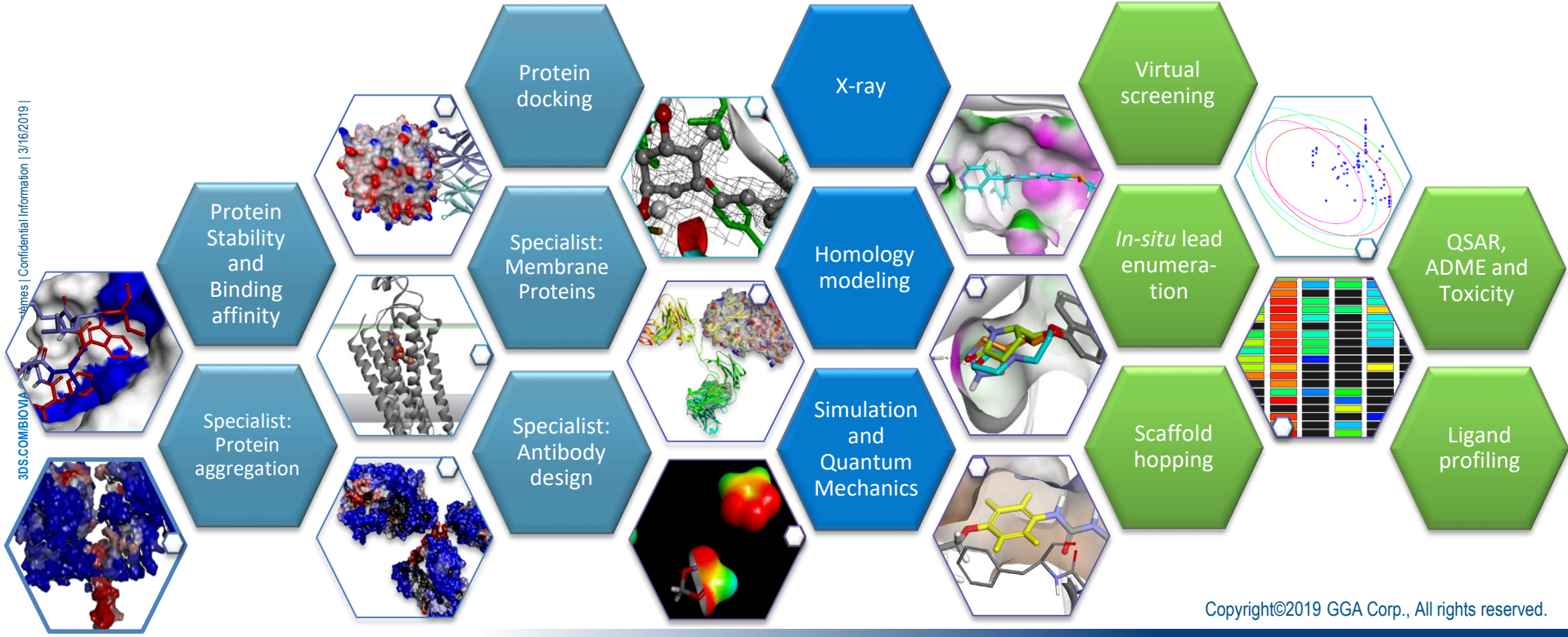
© 3DS_Document_2013
3DS.CO

What's Discovery Studio

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Discovery Studio: A Comprehensive Portfolio



3DS.COM/BIOVIA | ...mes | Confidential Information | 3/16/2019 |

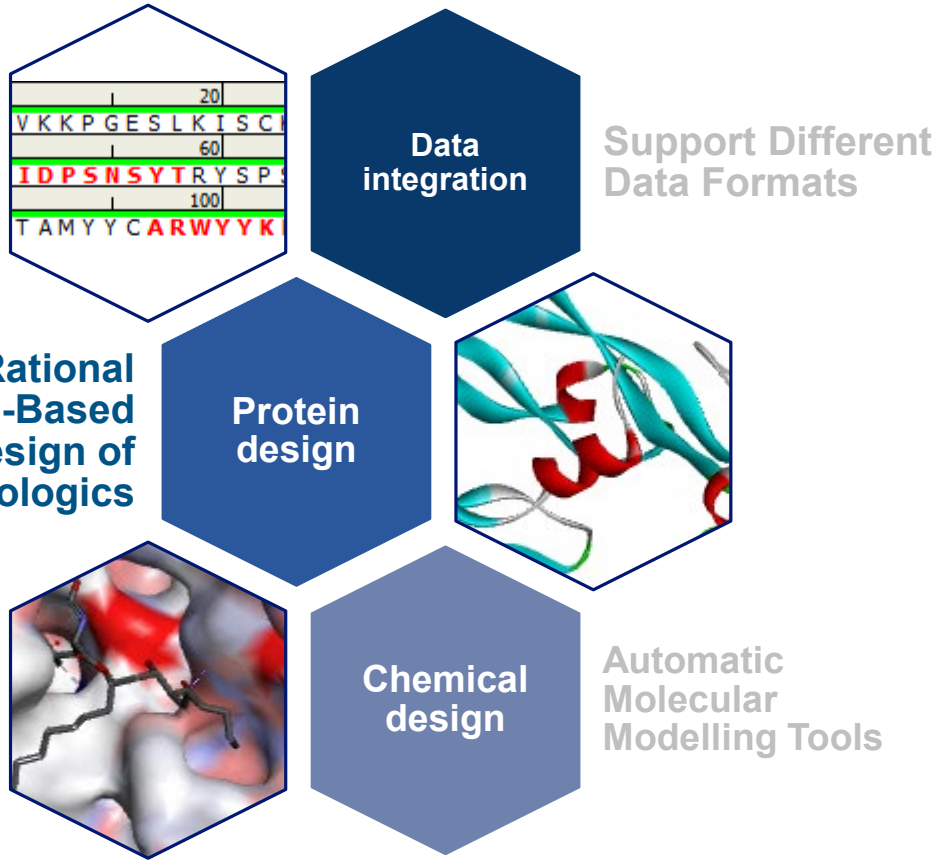
Copyright©2019 GGA Corp., All rights reserved.

Protein design tool includes:

1. Advanced homology modeling tools
2. Protein docking and epitope detection
3. pH-based protonation
4. Protein aggregation
5. Virtual mutation
6. …

Benefit

1. Increase the efficiency in sample preparation stage
2. Epitope prediction
3. Optimize protein

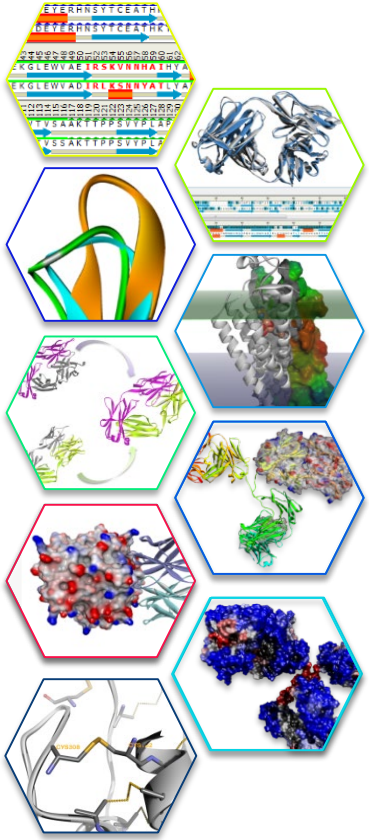


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Science and Functionality Highlights

- Template identification
- Sequence alignment and analysis
- **Advanced homology modeling tools**
- Rigorous model verification
- Implicit or explicit solvent-based MD simulations
- Specialist tools for Antibody modeling
- **Antibody loop modeling and refinement**
- **Protein docking and epitope detection**
- **Optimizing protein-protein binding affinity**
- Predicting and guiding protein stability
- **Managing aggregation risk**
- Disulfide bridge prediction
- **Humanization**

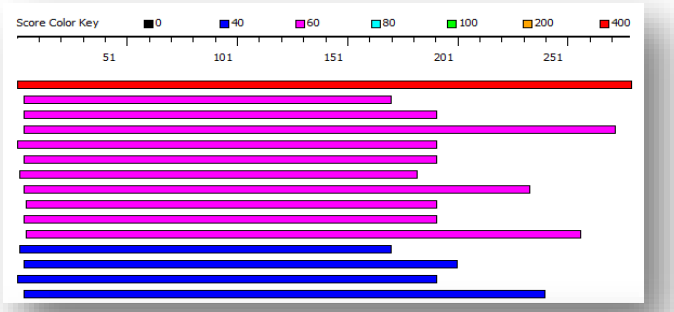


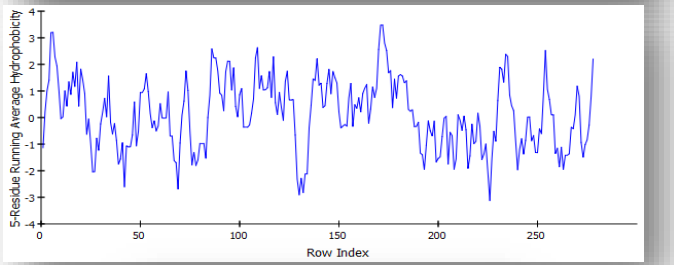
Copyright©2019 GGA Corp., All rights reserved.

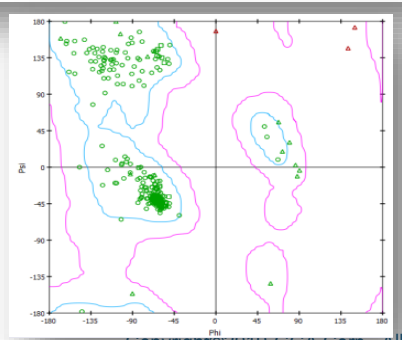
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Template Identification

- Search the PDB using Blast or PSI-BLAST
  - Identify optimal templates for the target macromolecule
- Generate protein reports
  - Spot potential problems with a retrieved template
- Refine template selection
  - Specify and filter by species





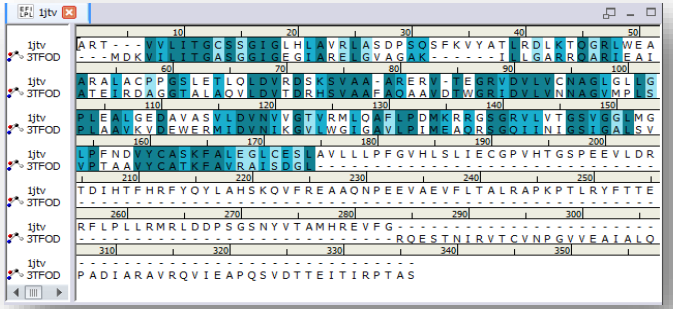


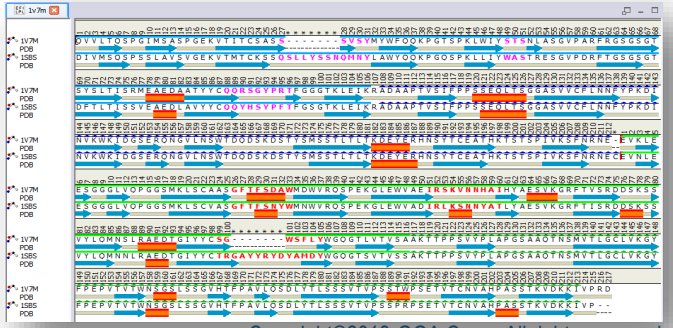
Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Sequence Alignment and Analysis

- Align sequences quickly and accurately to templates
  - Use structure alignment to align template sequences
  - Align123*: Align model sequences to the structure profile
  - Optionally, include secondary structure matching when aligning sequences using Align123
  - Optionally, predict helices in transmembrane proteins from sequence‡
- Identify sequence motifs associated with possible post-translational
  - modification sites in biotherapeutics
  - -
  - Additionally, calculate biophysical properties, such as
  - isoelectric points, molecular charge, molar extinction
  - coefficient, hydropathy and Antigenic sites





* Marti-Renom M.A., Madhusudhan M.S., Sali, A. *Protein Science,* **2004**, **13**, 1071-1087

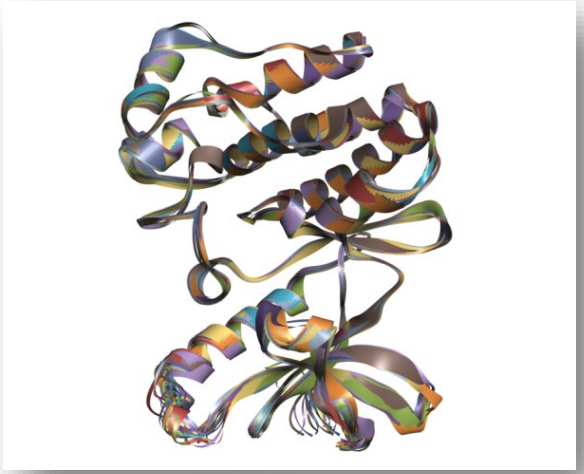
‡ Sonnhammer E. L., von Heijne G., Krogh A. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **1998**, 6, 175

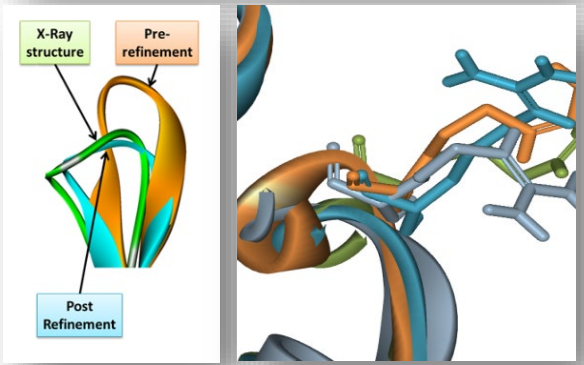
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Advanced Homology Modeling Tools

- Use the industry-standard MODELER†
  - Build homology models of target proteins automatically
  - Use LOOPER* and CHARMM to refine loops
  - Optimize side-chain conformation using CHARMm simulations





† Webb B., Sali A., *et al*, *Current Protocols in Bioinformatics, John Wiley & Sons, Inc.*, **2006**, Supplement 15, 5.6.1-5.6.30

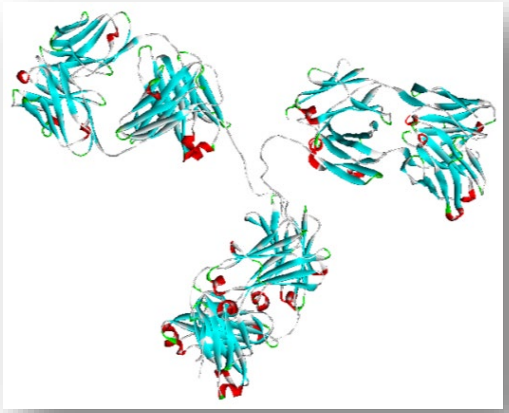
* Spassov, V.Z., Flook, P.K., Yan, L. *Prot. Eng., Design & Selection*, **2008**, 21, 91-100

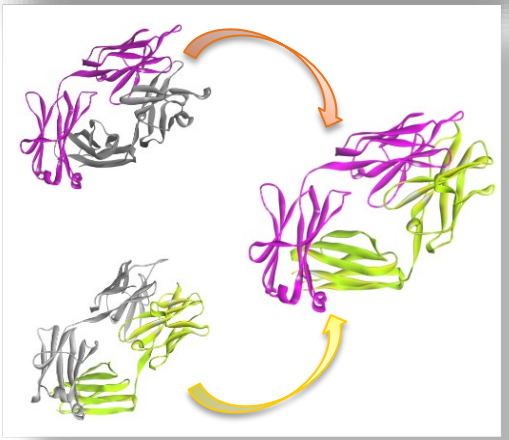
Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Specialist Tools for Antibody Modeling

- Best in class* mAb homology Modeling toolkit
  - Full-length IgG1 structure models
  - Fab or Fv domain (VHL) Framework models
  - VH and VL Template structure identification
  - Accurately model CDR loops
  - Annotate using IMGT, Chothia, Honegger or Kabat
  - Model refinements and simulations





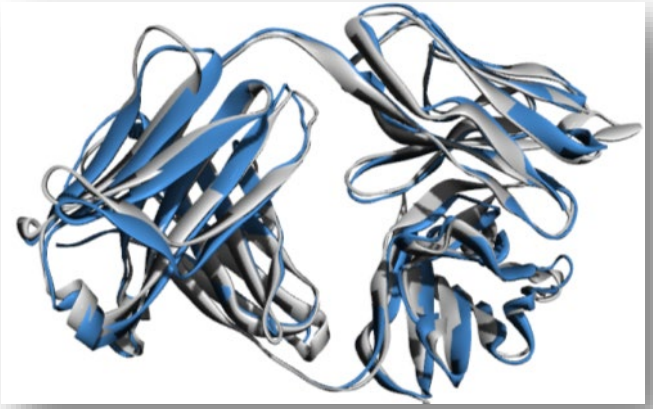
* Almagro J.C., *et al*, Proteins, **2011**, 79(11), 3050-3066 DOI: 10.1002/prot.23130

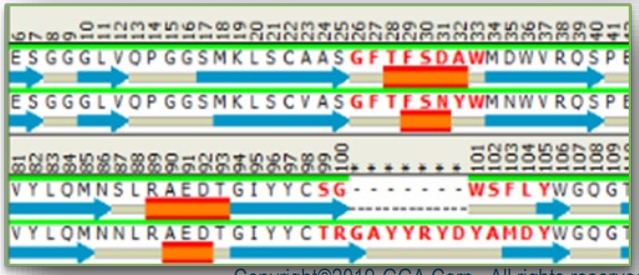
Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Specialist Tools for Antibody Modeling

- Identify Templates in PDB or curated PDB Antibody databases
  - Automatically identify variable and constant domains using HMM
  - Report CDR loop and number using: IMGT, Chothia, Kabat, Honegger
  - Search and identify optimal templates for each chain, or domain
  - Ability to exclude CDR regions in template search
  - Filter by organism (human, mouse, etc.)



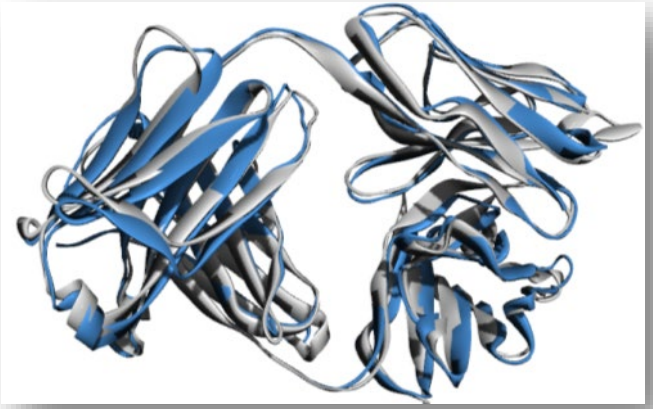


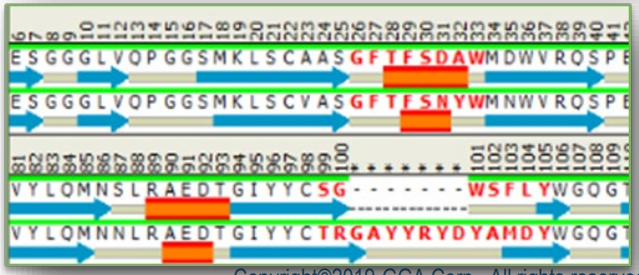
Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Specialist Tools for Antibody Modeling

- Quickly and accurately align model sequence with templates
  - Using multiple structure alignment and multiple sequence alignment algorithms
  - Using residue numbering including IMGT, Chothia, Kabat, Honegger
  - Simultaneously, but independently perform alignments on either light or heavy chains



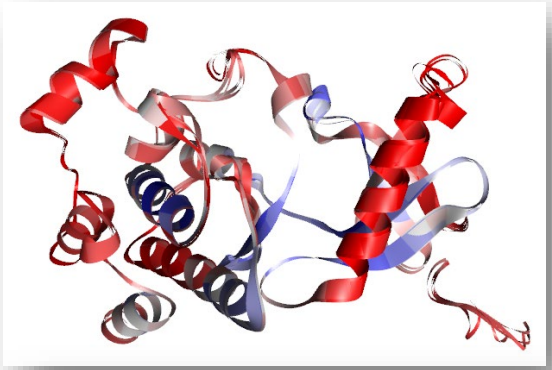


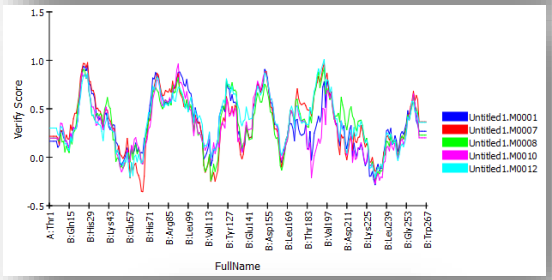
Copyright©2019 GGA Corp., All rights reserved.

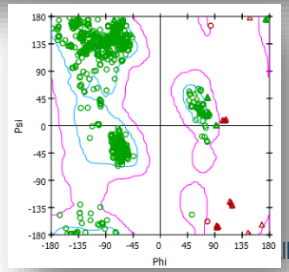
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Rigorous Model Verification

- Rigorously validate homology models
  - Verify Protein (Profiles 3D):
    - Evaluate the fitness of a sequence to its 3D environment
  - Verify Protein (MODELER):
    - Select the best structure from a collection of protein molecules
  - Ramachandran plots:
    - Verify *Phi* and *Psi* angle distributions





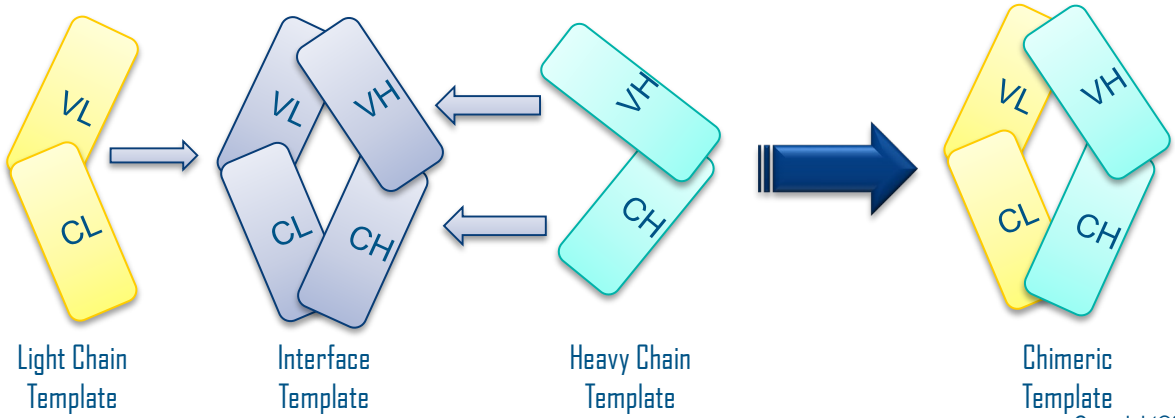


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

l rights reserved.

Specialist Tools for Antibody Modeling

- Use MODELER to build homology models of antibodies:
  - **Full Length:** Build models from IgG1 or IgG2 structures
    - Developed in collaboration with a large Pharma customer
  - **Framework:** Specify different templates for heavy chain and light chain respectively
    - Interface template to automatically orient chains

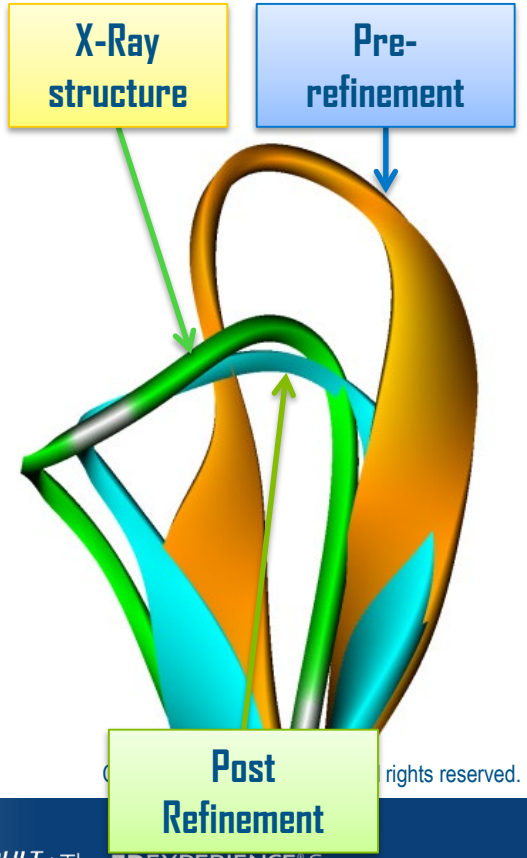


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Antibody Loop Modeling and Refinement

- Automatically identify the CDR loops using HMM
- Find the best template for each loop
  - Curated antibody database
  - Filter results by canonical structure type
- Build models based on loop templates
  - **Manual Loop Grafting:** Copy loops from a template structure
  - **Ab Initio Loop Refinement:** Use LOOPER* to systematically search and optimize CDR loops



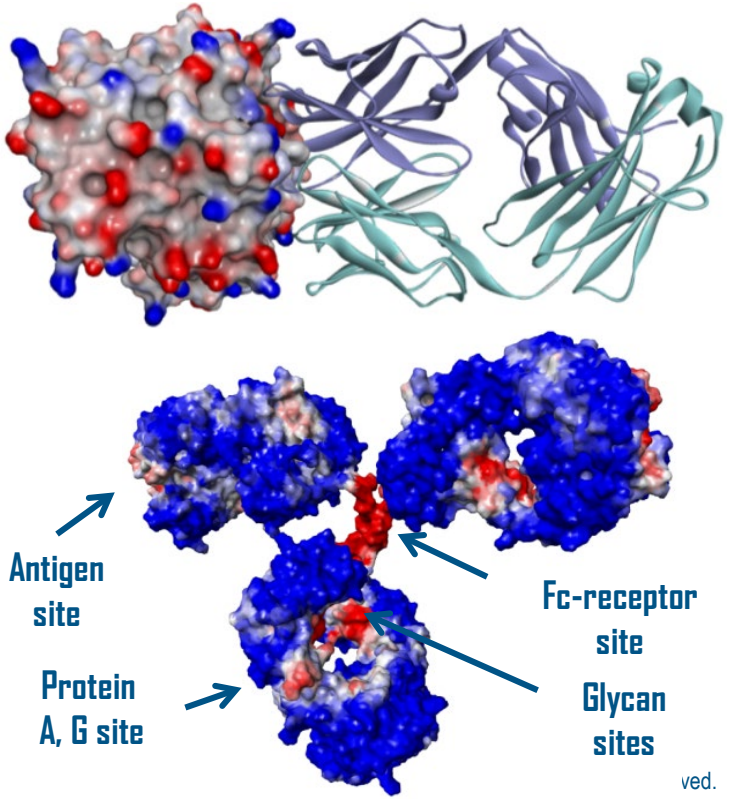
* Spassov, V.Z., Flook, P.K., Yan, L. Prot. Eng., Design & Selection, **2008**, 21, 91-100

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

© rights reserved.

Protein Docking and Epitope Detection

- Predict Antigen-Antibody binding:
  - **ZDOCK*†:** Comprehensively search protein-protein interaction patterns and output possible docking poses
  - **ZRANK‡:** Quickly and accurately refine ZDOCK rankings
- Additionally, use **AggMap#** to quickly identify putative protein-protein interaction sites

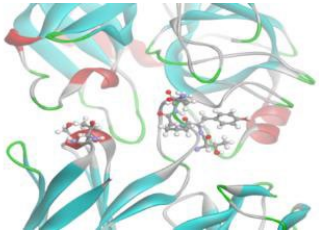


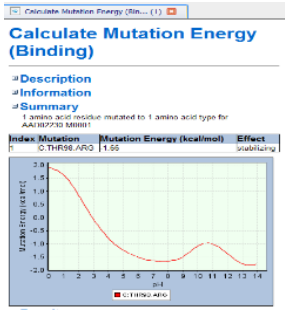
* Pierce B., Weng Z. Proteins, **2008**, 72(1), 270-279
† Chen R., Weng Z. Proteins, **2003**, 52, 80-87.
‡ Pierce B, Weng Z. Proteins, **2007**, 67(4), 1078-1086
# Chennamsetty N., Voynov V., Kayser V., Helk B., and Trout B.L., Proteins, **2011**, 79, 888–897

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |
ved.

Virtual Mutagenesis: Predict mutation effect for both binding and stability

- Simple, consistent and flexible user interface made it easy for:
  - Virtual alanine scanning (or any type of residue scanning) X → Ala
  - Virtual saturated mutagenesis X → [A C D … V W Y]
- Automatically generate multiple mutations
- pH-dependent mutations to take into account of solution pH and protein ionization
  - wild type and mutant structures with the proper protonation for studying their electrostatics properties
  - pH dependent binding/stability for wild type and for mutants
  - pH dependent mutation energies
- Consider ionic strength of solution
- Consider temperature



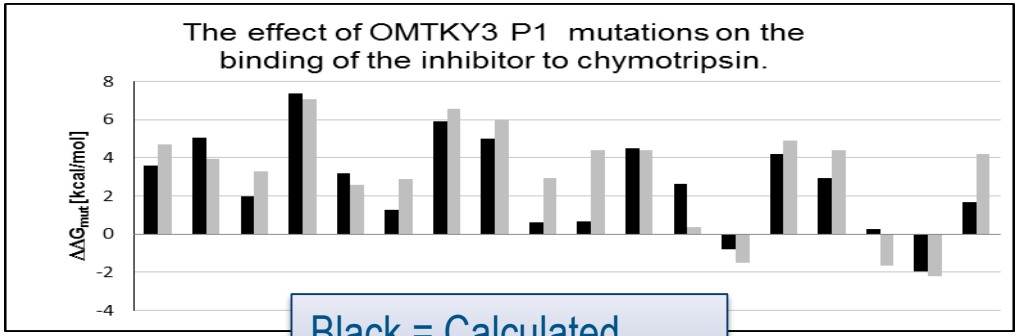


Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Optimizing Protein-Protein Binding Affinity

- Evaluate the effect of point mutations on binding to molecular partners
  - Perform ‘Ala-scanning’: X ➡ Ala
  - Or, full residue type scanning: X ➡ [A C D … V W Y]
- Calculate the effect at a specific temperature or pH
  - Based on published* experimental $\Delta\Delta G_{bind}$ energies for 380 single point mutations across 19 proteins



* Spassov, V.Z., Flook, P.K., Yan, L. Prot. Eng., Design & Selection, **2008**, 21, 91-100

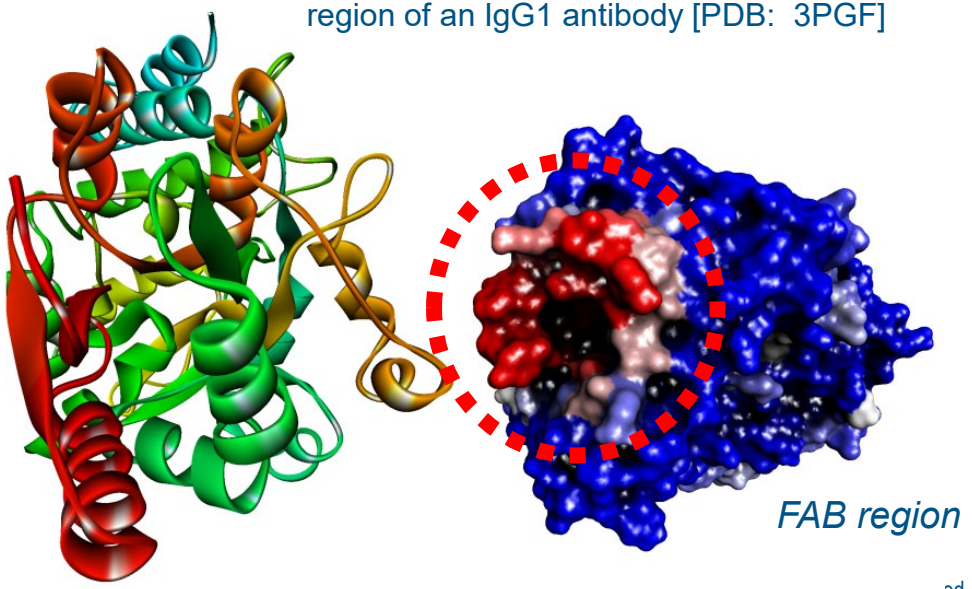
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

9 GGA Corp., All rights reserved.

Spatial Aggregation Propensity – Other Uses

- Protein aggregation hotspots are useful for other applications beyond Antibody formulations*
  - E.g., prediction of protein binding regions

E.g., Antigen binding domain on the FAB region of an IgG1 antibody [PDB: 3PGF]



* Chennamsetty N., Voynov V., Kayser V., Helk B., and Trout B.L., Proteins, **2011**, 79, 888–897

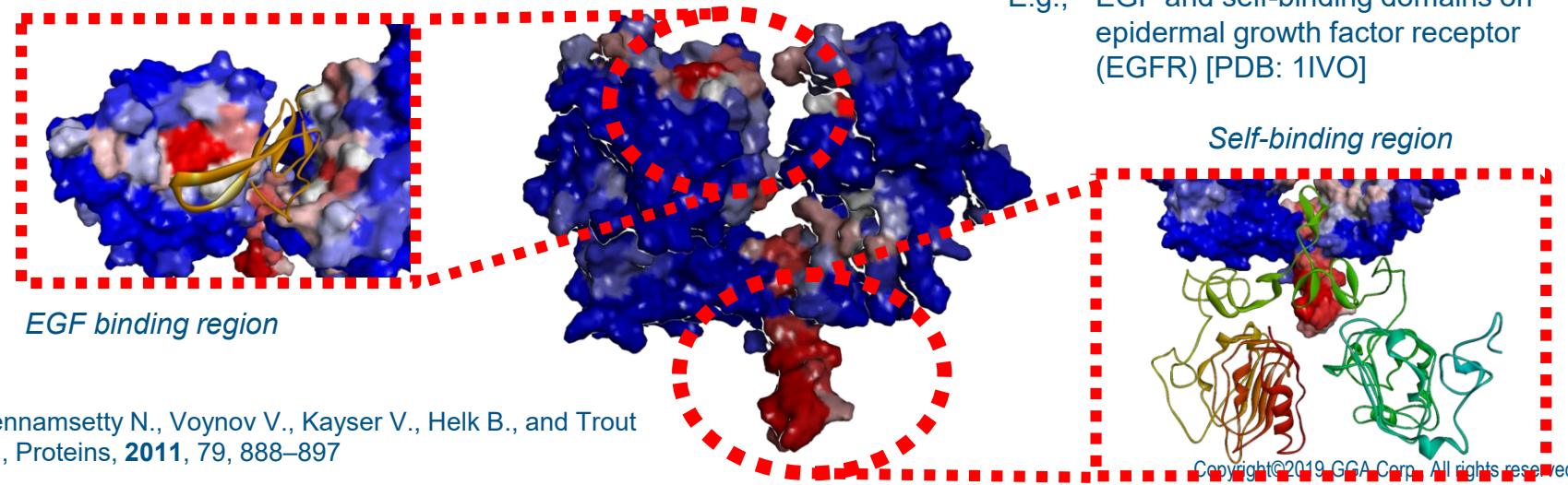
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

ed.

Spatial Aggregation Propensity – Other Uses

- Protein aggregation hotspots are useful for other applications beyond Antibody formulations*
  - E.g., prediction of protein binding regions

E.g., EGF and self-binding domains on epidermal growth factor receptor (EGFR) [PDB: 1IVO]



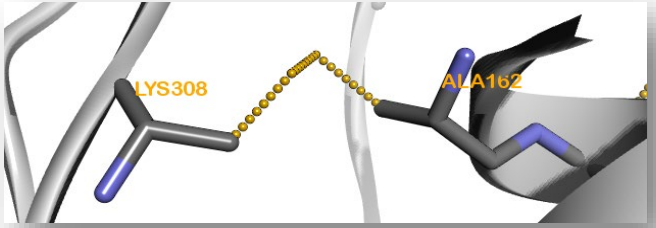
* Chennamsetty N., Voynov V., Kayser V., Helk B., and Trout B.L., Proteins, **2011**, 79, 888–897

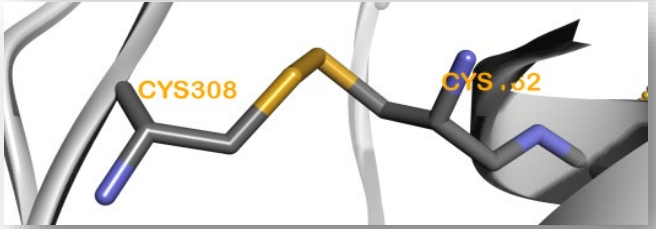
Copyright©2019 GGA Corp. All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Disulfide Bridge Prediction

- Improve stability in a protein, by predicting disulfide bridge sites
  - Validated using *ca.* 1,500 structures with genuine disulfide bridges
- Disulfide bonds assessed for both good stereochemistry and factors such as:
  - Steric clashes, thermal mobility and residue depth





Predicted disulfide bridges for molecule 1TCA

Predicted Disulfide Bond	Original Residues	Score	Energy	Bad Contacts	Thermal Mobility	Sequence Separation	Residue Depth	Volume Change	Environment	Chi3
CYS239:SG - CYS252:SG	A:SER239, A:ASP252	97.52	1.24	0	N/A	13	3.88	-5.4	Helix-Helix	-91.8
CYS162:SG - CYS308:SG	A:ALA162, A:LYS308	96.99	1.51	0	N/A	146	4.63	-52.5	Coil-Helix	-86.4
CYS189:SG - CYS224:SG	A:ILE189, A:HIS224	96.83	1.58	0	N/A	35	5.60	-115.9	Coil-Turn	116.8
CYS34:SG - CYS61:SG	A:ILE34, A:TYR61	95.80	2.10	0	N/A	27	6.93	-153.6	Coil-Sheet	129.8
CYS115:SG - CYS176:SG	A:LEU115, A:ILE176	94.57	2.71	0	N/A	61	5.80	-120.9	Coil-Helix	134.7
CYS50:SG - CYS273:SG	A:SER50, A:ALA273	93.10	3.45	0	N/A	223	3.85	22.5	Helix-Turn	88.4
CYS156:SG - CYS163:SG	A:GLN156, A:LEU163	92.97	3.51	0	N/A	7	7.78	-107.3	Helix-Helix	119.5
CYS180:SG - CYS240:SG	A:THR180, A:ALA240	92.75	3.62	0	N/A	60	8.17	-3.2	Helix-Sheet	-94.4
CYS183:SG - CYS202:SG	A:TYR183, A:SER202	92.59	3.70	0	N/A	19	9.66	-84.6	Sheet-Turn	-95.8
CYS103:SG - CYS108:SG	A:THR103, A:GLY108	92.43	3.78	0	N/A	5	15.82	22.5	Helix-Sheet	110.7
CYS262:SG -	A:PRO262,	92.20	3.90	0	N/A	12	5.44	-6.0	Coil-Helix	-

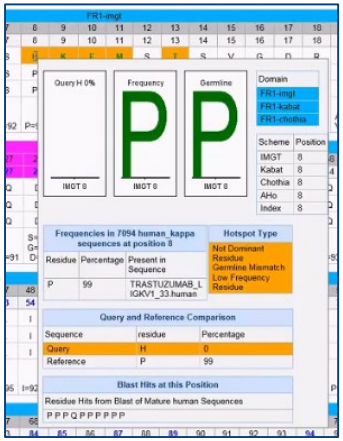
In testing, successfully reproduces the engineered position reported by Kim *et al* in lipase B [PDB: 1TCA] DOI: 10.1002/bit.24371

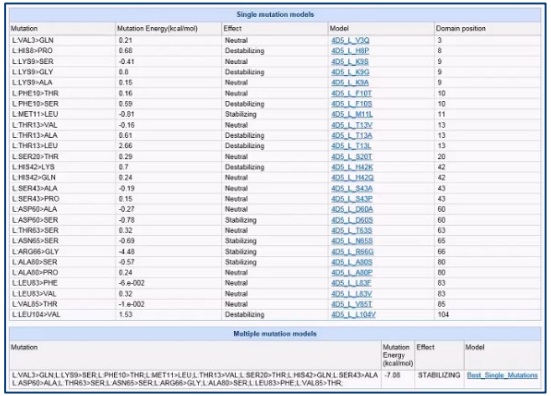
Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Antibody Humanization Workflow

- Search query Light and Heavy chain antibody sequences for humanizing mutations
  - Search against reference germline databases
  - Optionally, compare to a reference sequence
- Highlight residue humanization hotspots
  - Sites where residue occurs in less than 5% frequency
    and do not match germlines
  - See humanizing residues *via* ‘logo plots’
  - Optionally, exclude CDR and Vernier regions
- Predict most stabilizing humanized mutations
  - Either supply a pre-built antibody structure model
  - Or, use the Antibody Modeling Cascade to automatically build a 3D model

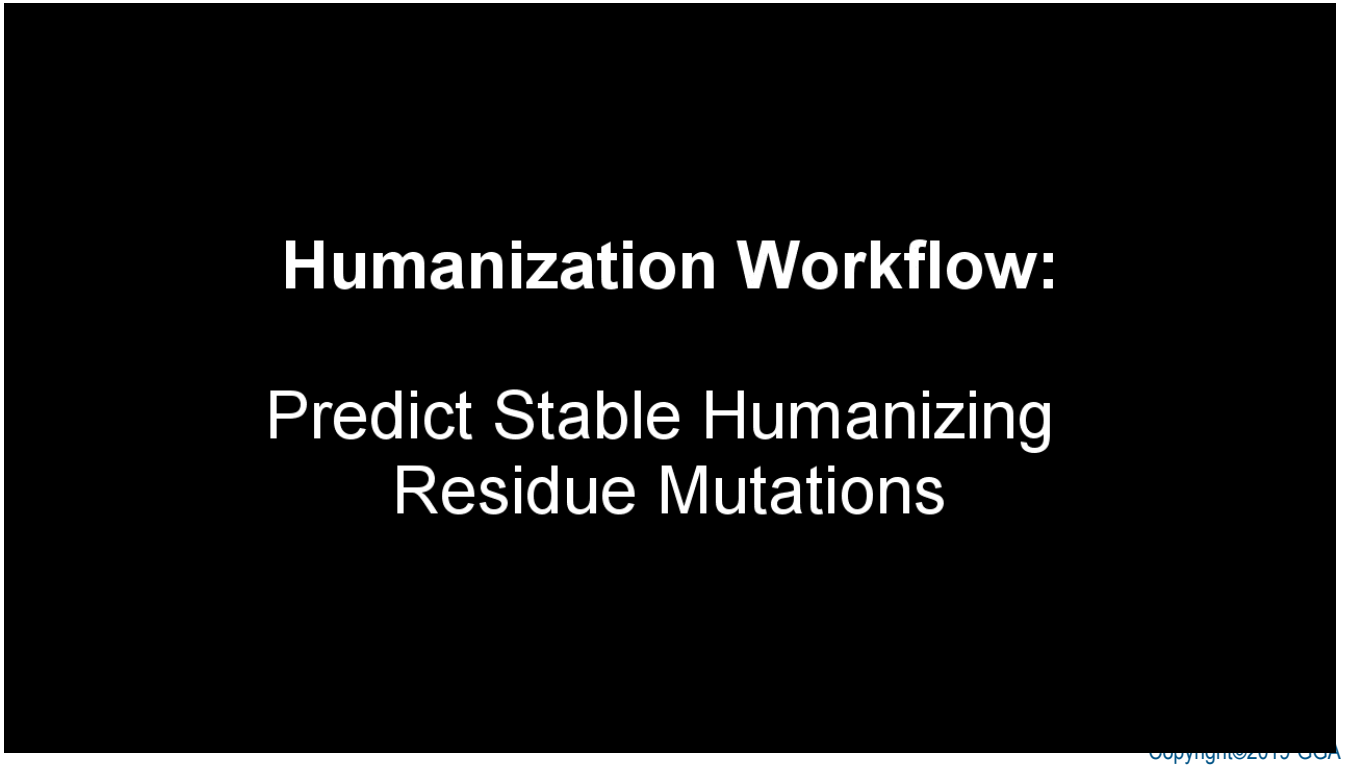




Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

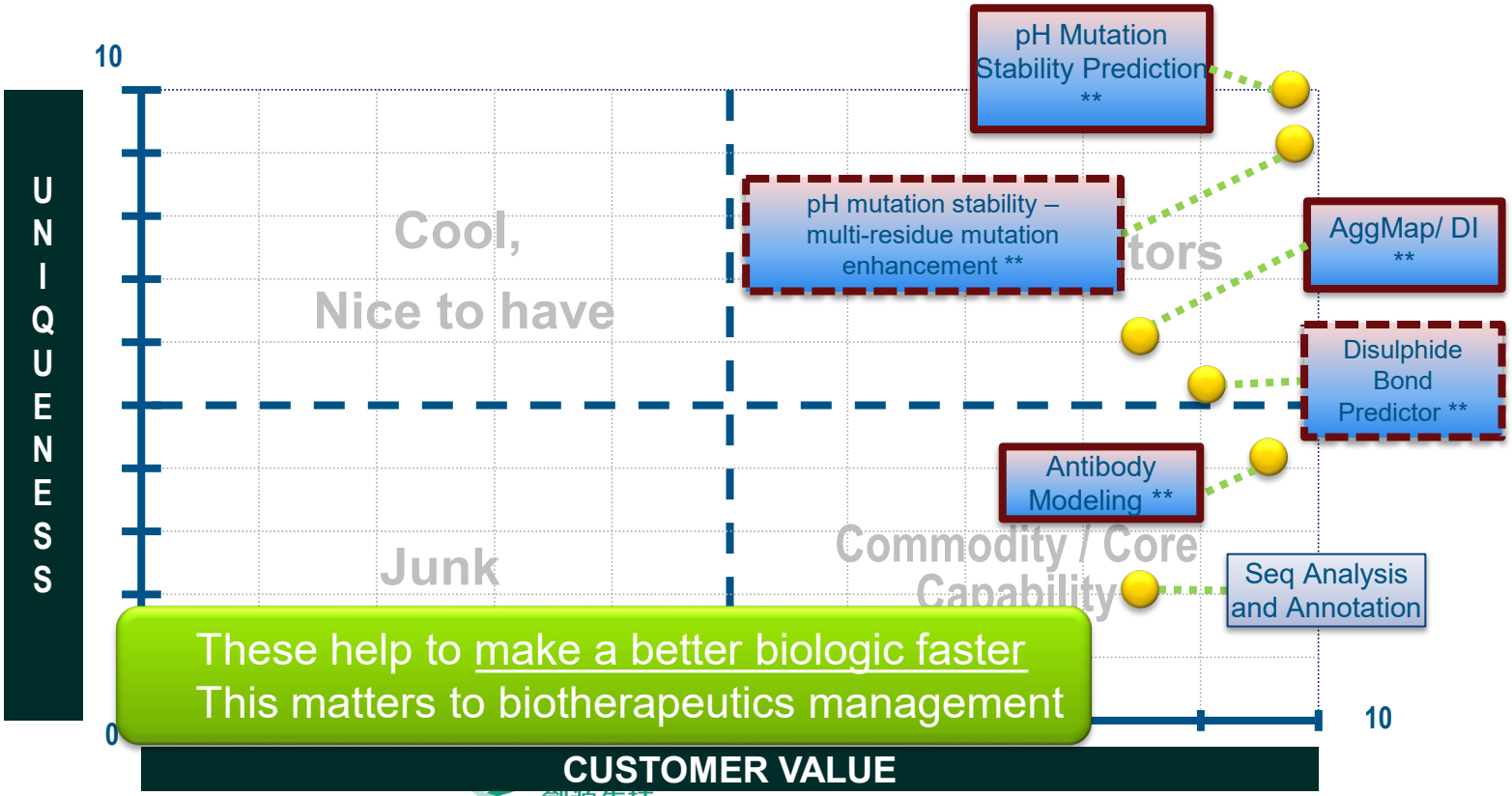
Antibody Humanization Workflow



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

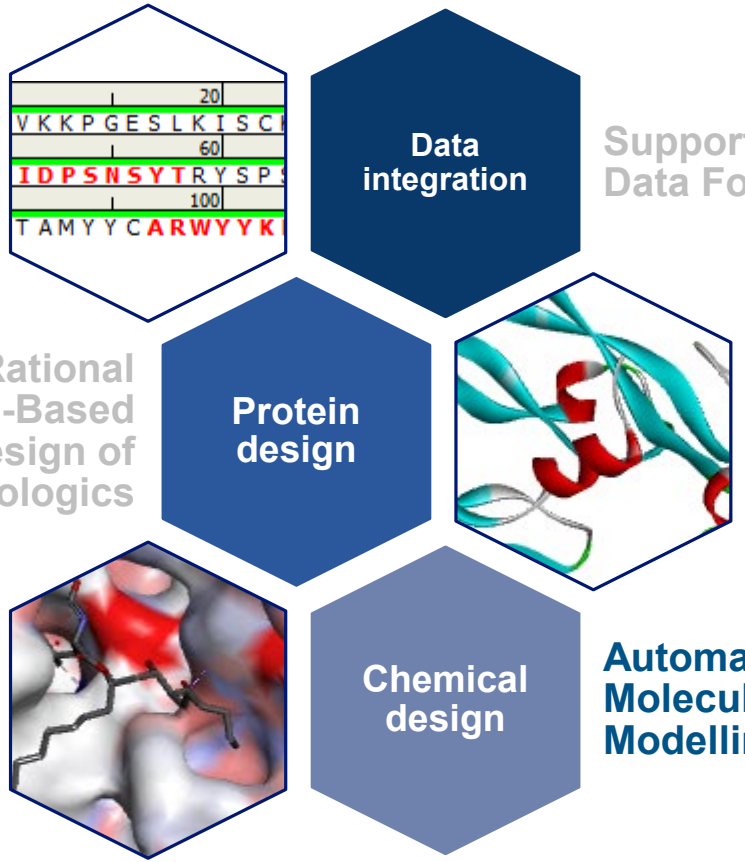
Solution Component Mapping Chart



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技 GGA

Copyright©2019 GGA Corp., All rights reserved.



Chemical design tool includes:

1. Structure-based design tools
2. Ligand-based design tools
3. Fragment-based design tools
4. Library design tools
5. ADMET and Toxicity prediction tools

Benefit

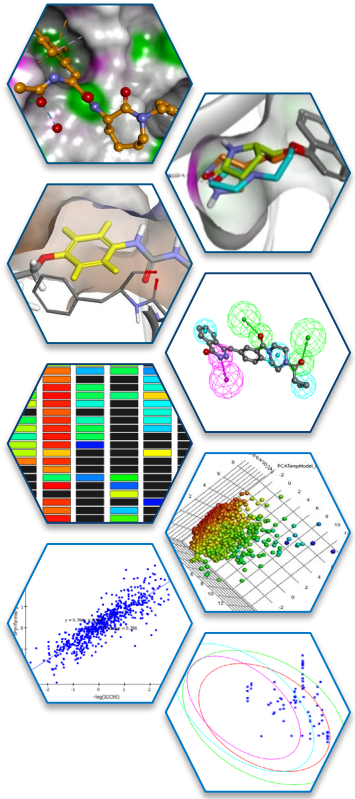
1. Explore the drug mechanism
2. Reduce the cost and failure rate
3. Virtual High Throughput Screen (vHTS)
4. Target fishing
5. Lead optimization

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Science and Functionality Highlights

- Structure-based design
  - Docking
  - Fragment-based design
  - Scoring/Prioritisation
  - Refinement/ Filtering/Analysis
- Pharmacophore modelling
  - Ligand-based
  - Structure-based
  - Ligand Profiling
- Library Design and Analysis
- Structure Activity Relationships
  - ADMET and Toxicology
  - QSAR and Machine Learning



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Rationale Small Drug Design

Scales

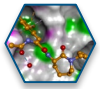
10^3~10^7 Molecules from Library

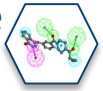
10^0~10^2 Molecules As Leads

10^2~10^3 Molecules as Novel Derivatives

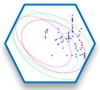
10^0~10^2 Molecules Pass ADMET and Toxicity Screening

Library Design and Analysis 

Structure-based design 

Pharmacophore modeling 

Fragment-based design 

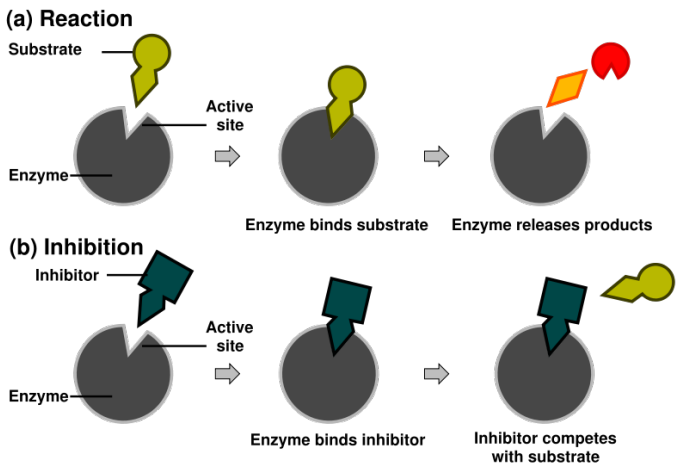
ADMET and Toxicity Analysis 

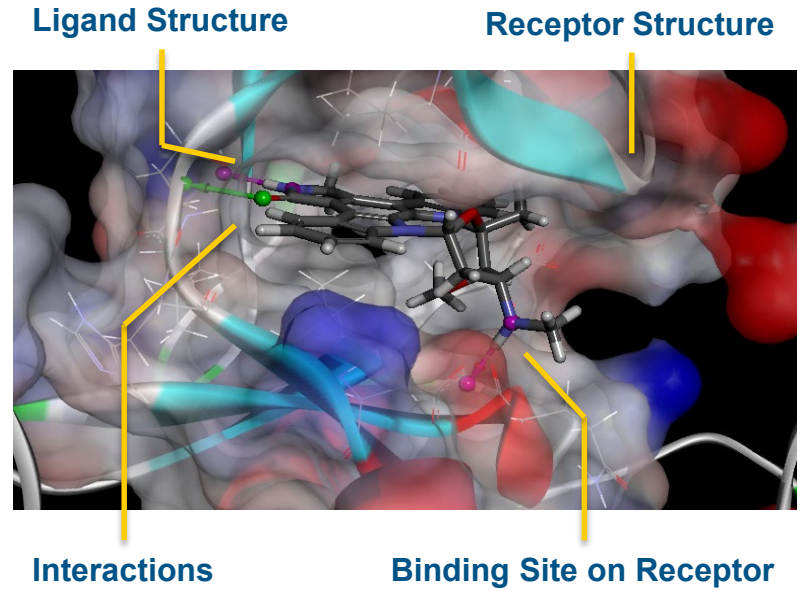
What if the hit has multiple targets? (Side effects)

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking & Scoring – Structure-Based Design



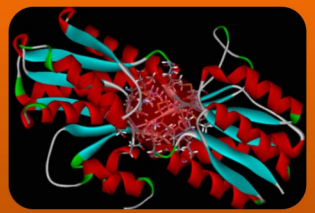


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

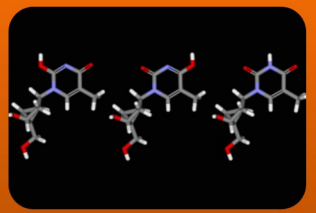
SBD: Input Preparation

Proteins



- Standardise atom names
- Insert missing atoms in residues
- Remove alternate conformations
- Insert missing loops
- Optimize short & medium size loops
- Calculate pK and protonate

Ligands



- Add hydrogens
- Calculate 3D coordinates
- Enumerate ionization states
- Ionize functional groups
- Generate tautomers and isomers
- Remove duplicates
- Fix bad valencies
- Standardize charges for common groups
- Retain largest fragment

Fragments



- Generate fragments using RECAP* rules
- Rule of Three filters

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

SBD: Fragment-Based Design Methods

Probing the Binding Site



Adding Fragments

Modifying Scaffolds

Alternative fragment-based methods available
- Fragment Based Pharmacophores

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

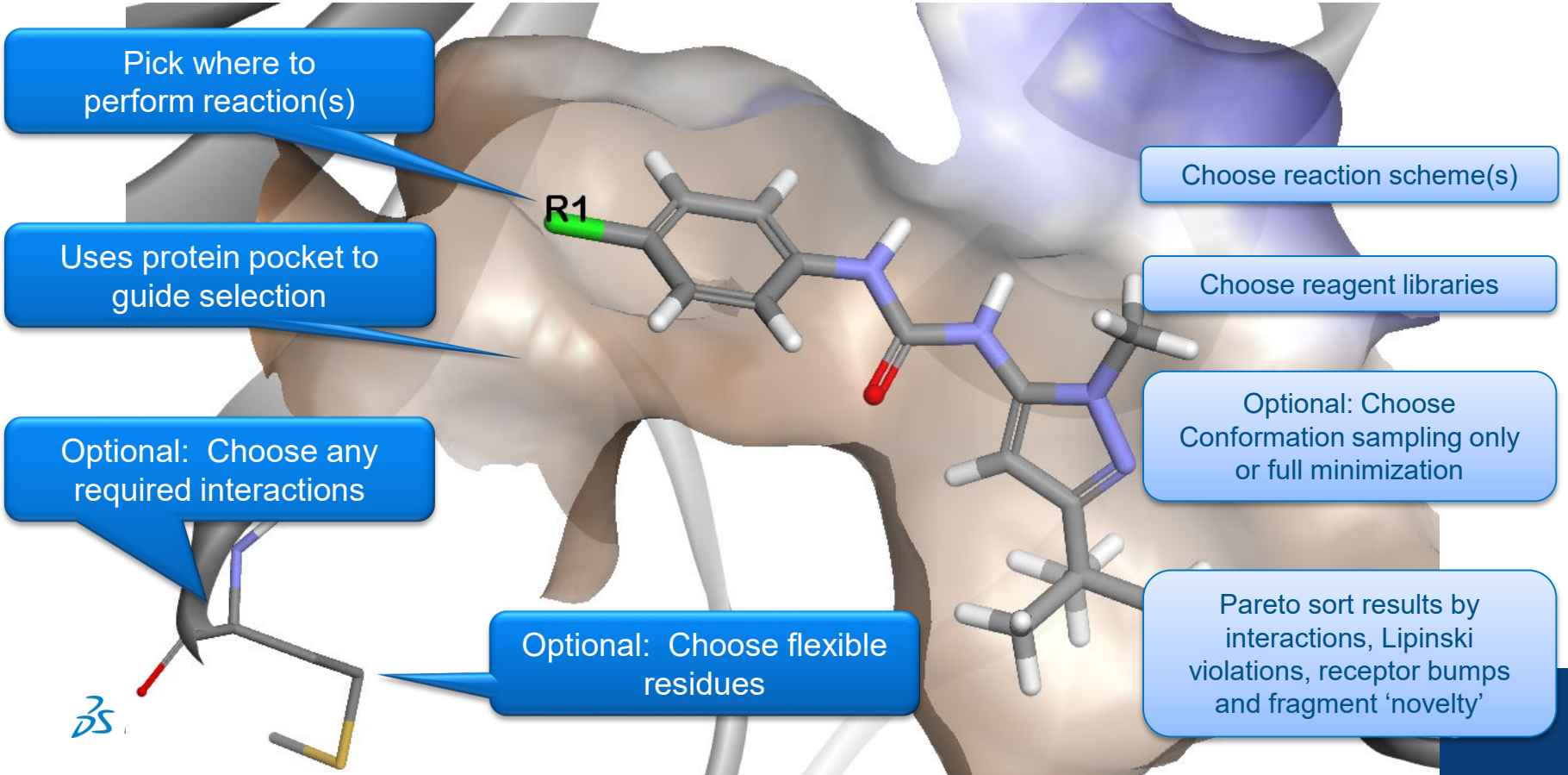
SBD: Fragment-Based Design Methods

- GROW
  - Reaction-based in situ ligand enumeration
  - E.g., Amide synthesis, Esterification, Hiyama, Kuyama, Negishi, Stille, Suzuki, Williamson Ether
  - Pre-filtered sets of reagents selected from ACD
- REPLACE
  - Fragment based in situ isostere replacement
  - E.g., scaffold-hopping, R-group replacement
  - Pre-filtered set of 1.5M fragments generated from SCD

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

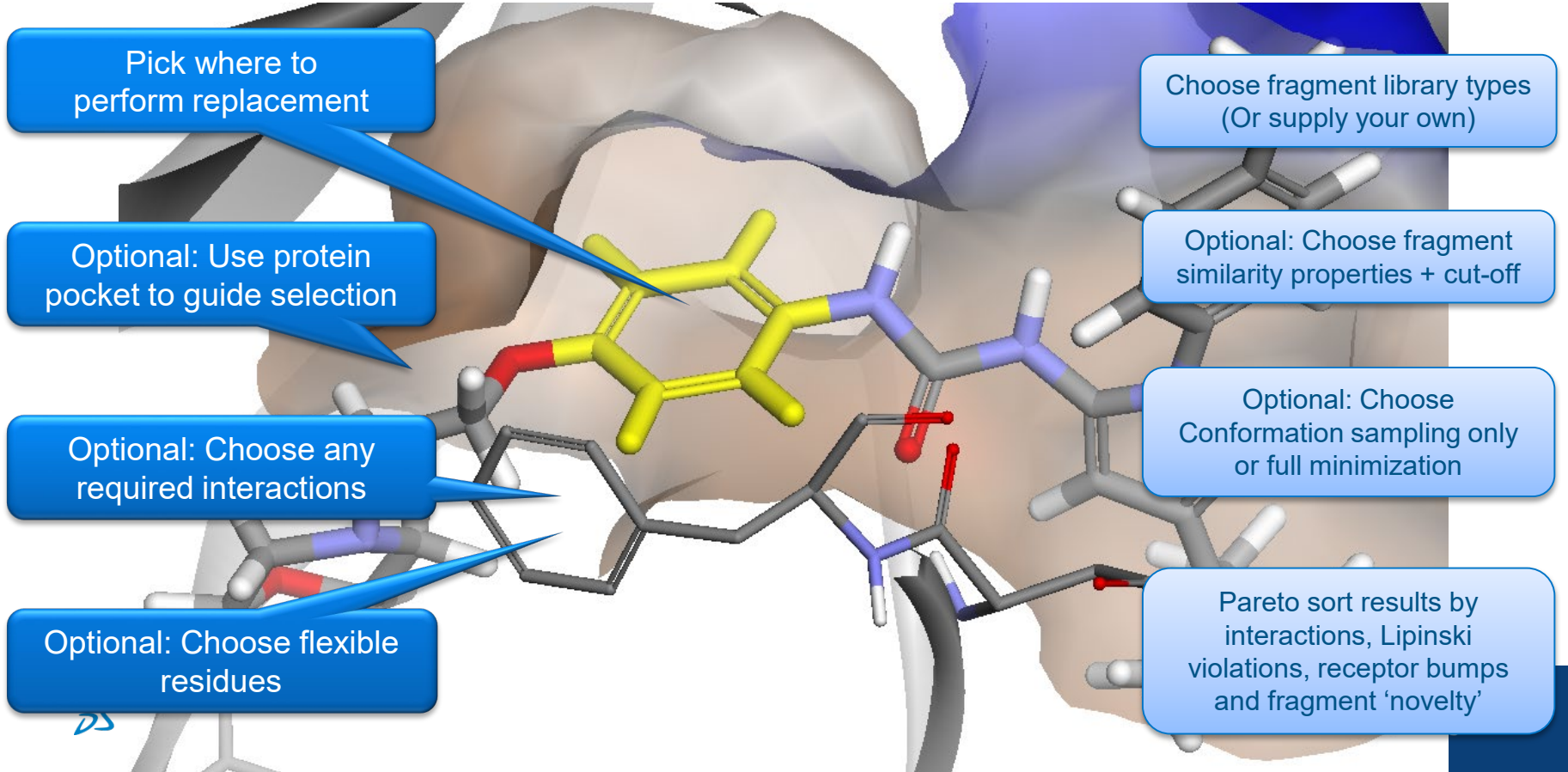
Copyright©2019 GGA Corp., All rights reserved.

GROW: Reaction-based *in-situ* Ligand Optimization



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 for MDOU and NYMU

REPLACE: Fragment Based *In-Situ* Substitution

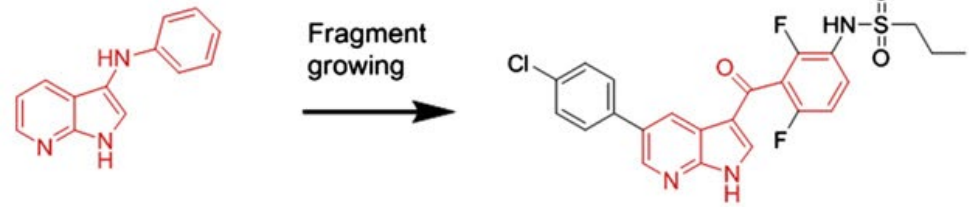


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Fragment-based design of the BRAF inhibitor vemurafenib.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

First fragment-based drug (Zelboraf) approved in 2011!



4

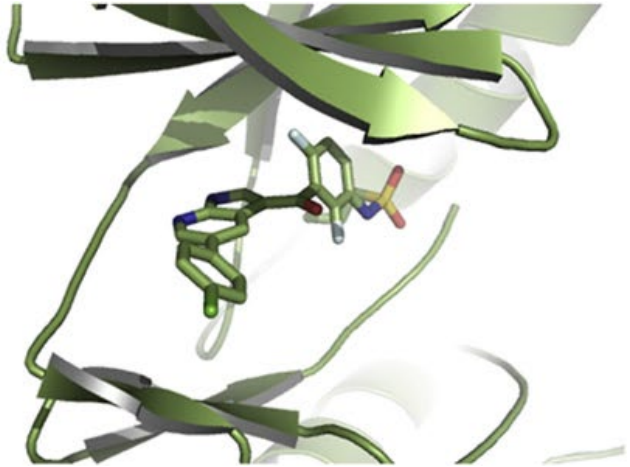
Unselective, weakly potent fragment hit (IC_{50} > 100μM)

PIM-1 IC_{50} ~ 100μM

5 (Vemurafenib)

BRAF (V600E) IC_{50} = 31nM
High degree of selectivity against other kinases

PIM-1 IC_{50} > 100μM

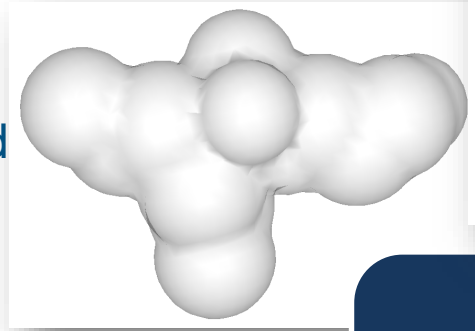


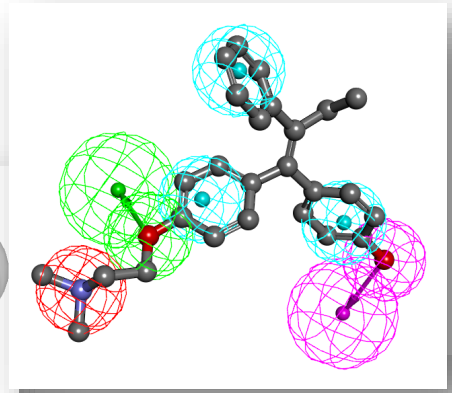
Swen Hoelder, Paul A. Clarke, Paul Workman, 2012

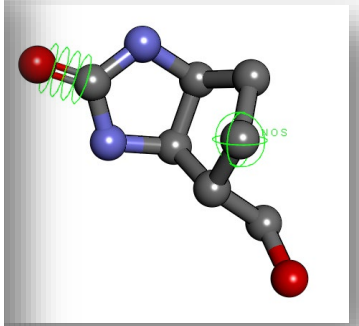
Copyright©2019 GGA Corp., All rights reserved.

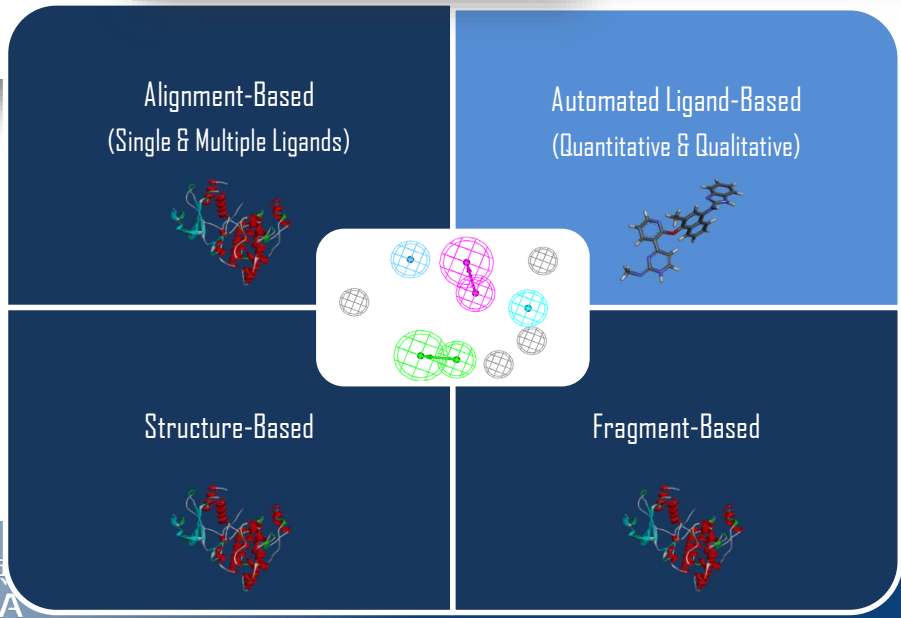
Pharmacophore Modelling

- 2D and 3D
  - SMARTS
  - Fragment-based
  - Feature-based
  - Shape
  - Combination
- Customisable pharmacophore features
- Automatic and manual generation
- Ligand-based and structure-based





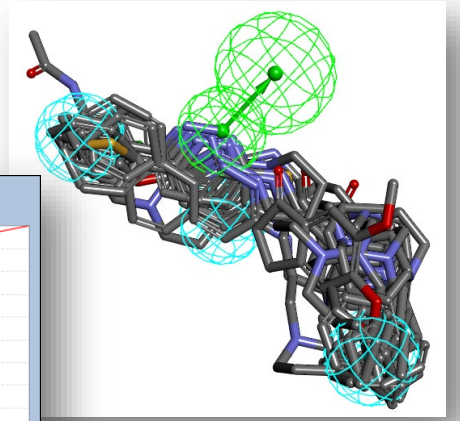


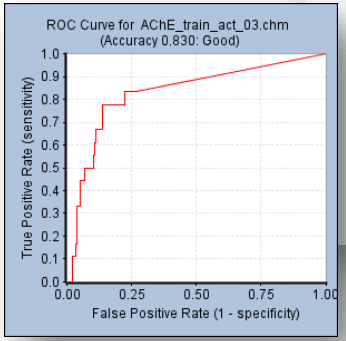


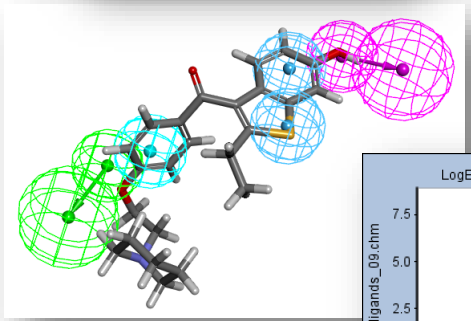
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

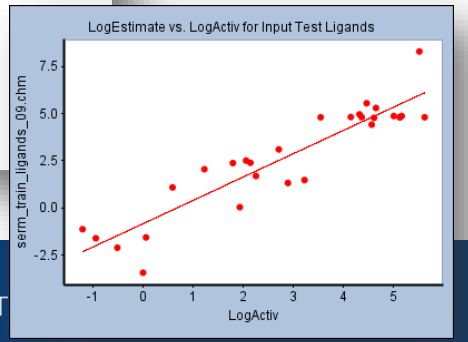
Ligand-Based Pharmacophores

- Automatic
  - Qualitative
    - Common features (HipHop / HipHopRefine)
    - Finds features shared by a set of similarly active ligands
  - Quantitative
    - SAR predicative (HypoGen / HypoGenRefine)
    - Finds features that relate to activity
  - Bioactive ligand conformation
- Manual
  - Alignment





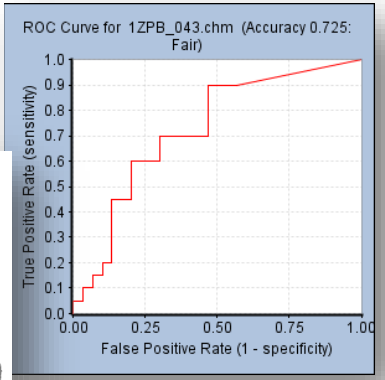


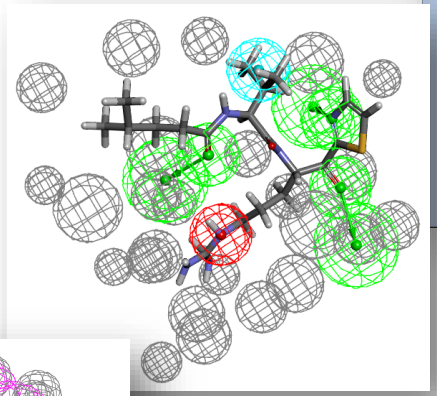


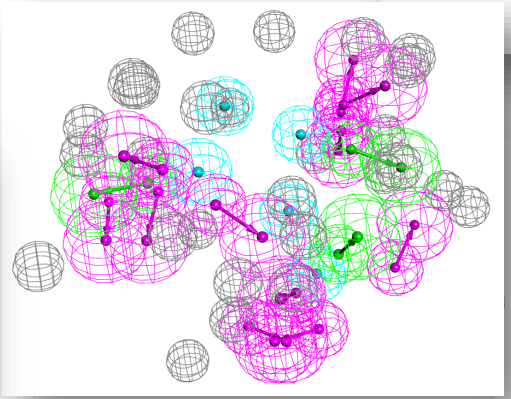
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

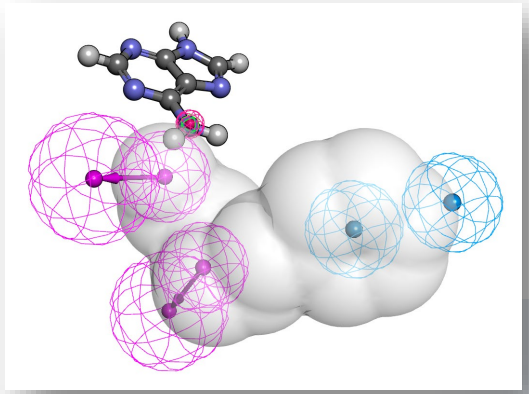
Structure-Based Pharmacophores

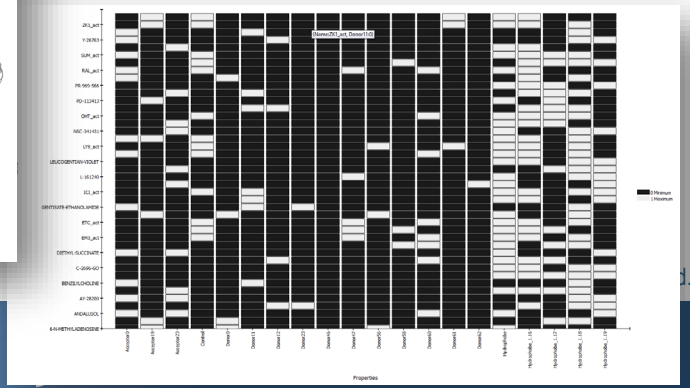
- Automatic
  - Receptor-Ligand complex
  - Interactions from a binding site
- Manual
  - Interaction map
  - Fragment-based









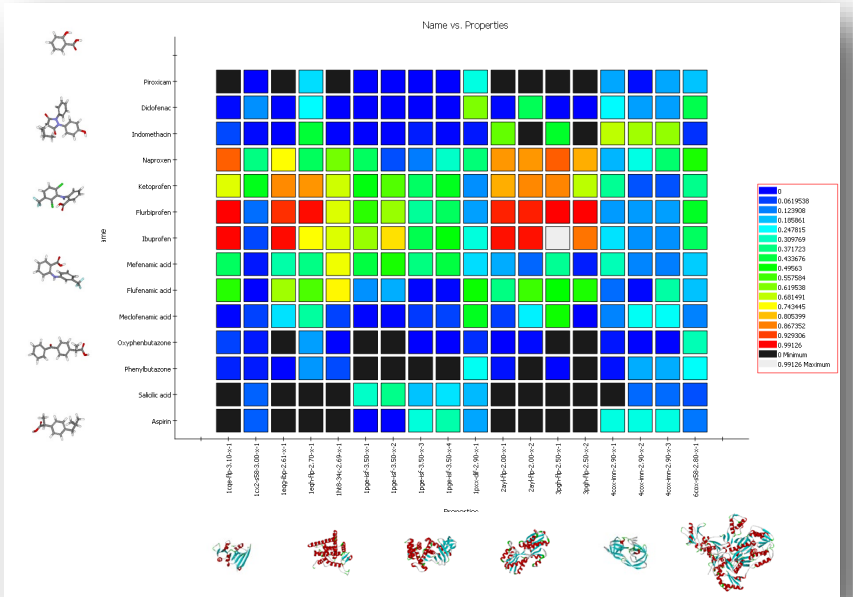


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技
GGA

Ligand Profiling

- Rapidly screen libraries of ligands against multiple pharmacophores
  - Predicting protein-drug off-targets (side effects)
  - Repositioning/repurposing existing drugs
  - In silico target fishing
- PharmaDB
  - Validated in collaboration with Prof. Rognan at University of Strasbourg*
  - Derived from the scPDB (http://bioinfo-pharma.u-strasbg.fr/scPDB)
  - 14031 validated models
  - Classified using Kyoto Encyclopedia of Genes and Genomes (KEGG)-BRITE



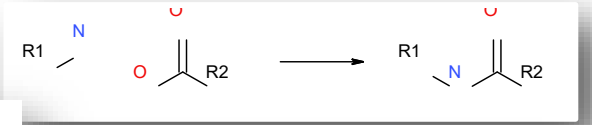
* Kellenberger *et al, J Chem Info Model*, **2006**, 46, 717-727

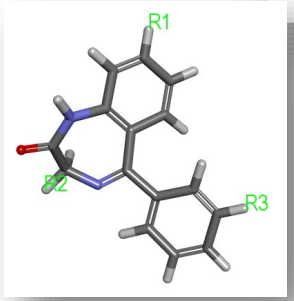
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

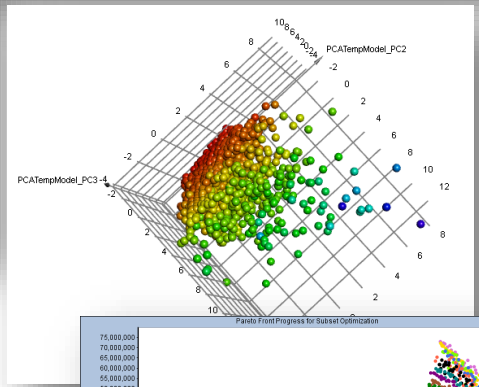
Copyright©2019 GGA Corp., All rights reserved.

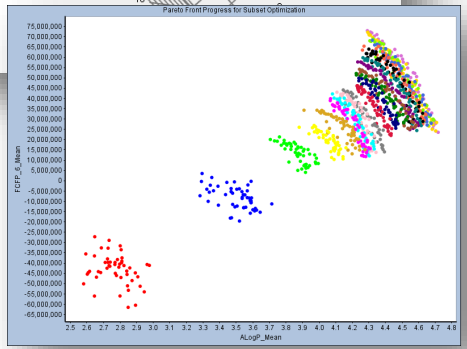
Library Design and Analysis

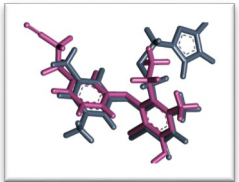
- Library enumeration
  - Reaction-based
  - Markush-based
- Diversity selection
- Clustering
- Similarity selection
- Multi-objective pareto optimization
- Novel ligand generation
  - BREED
- DS Application Edition (Pipeline Pilot)

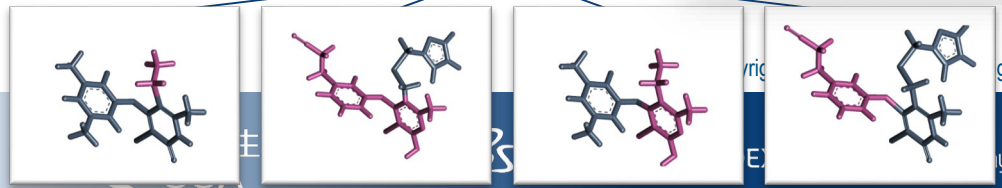










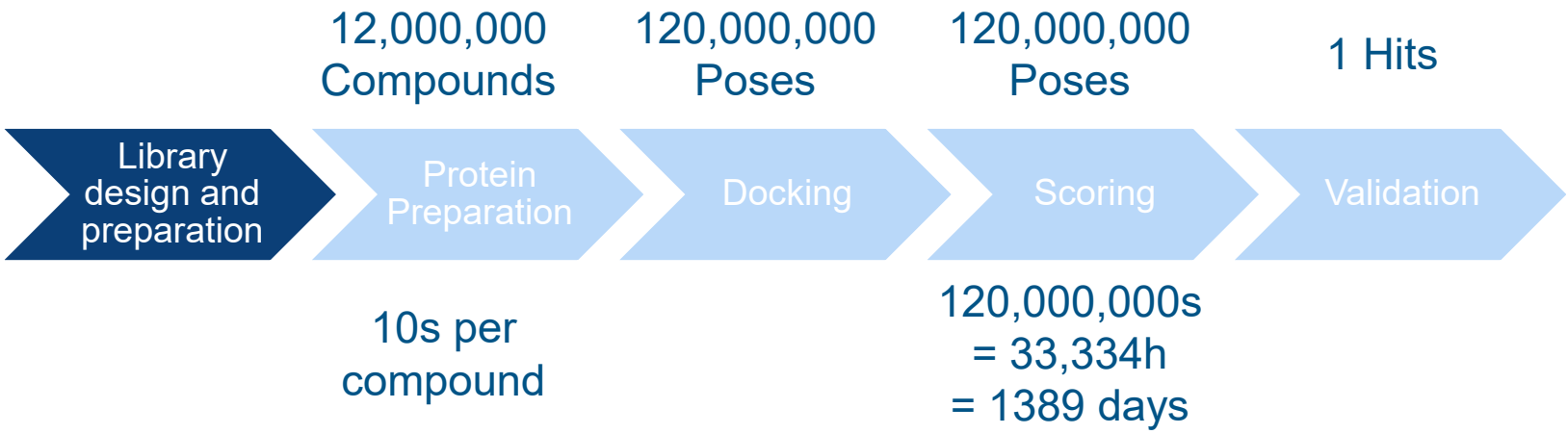


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

ghts reserved.

BIOVIA

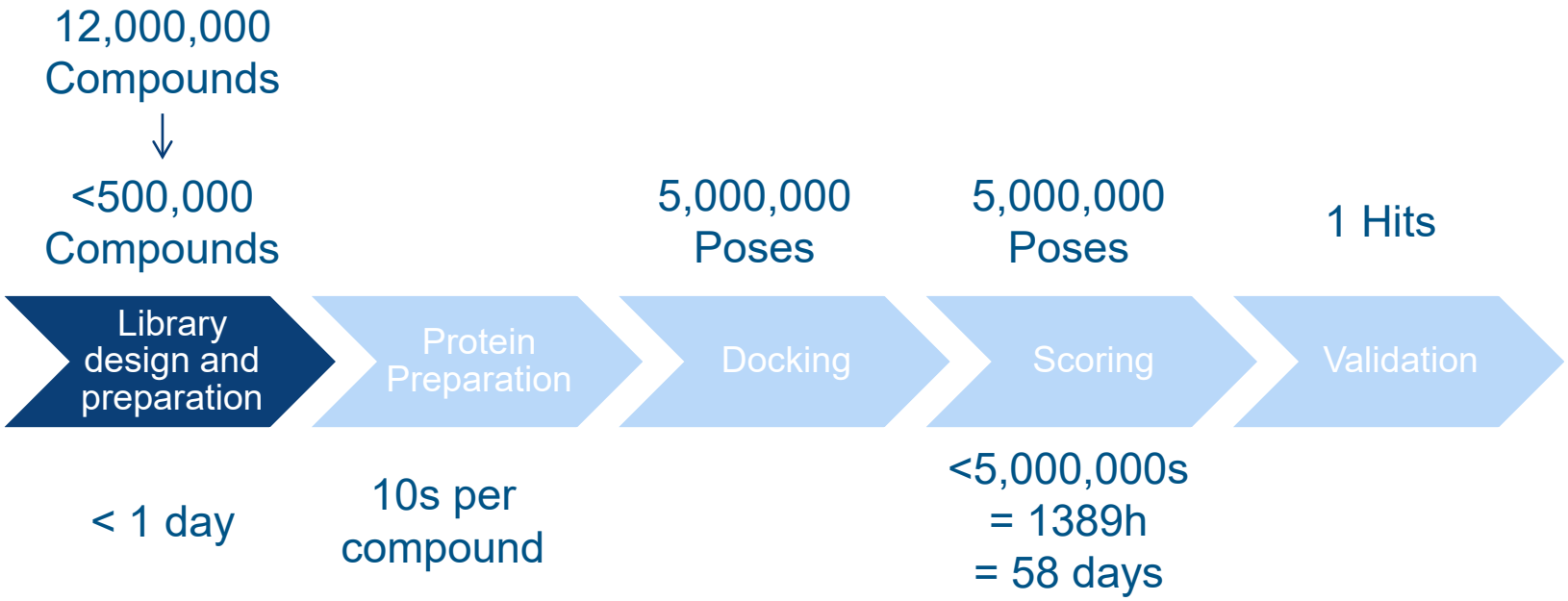
Workflow of Virtual High Throughput Screening



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Workflow of Virtual High Throughput Screening



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Compound databases

- BIOVIA database
  - **A**vailable **C**hemicals **D**irectory – 12,138,856 unique compounds
  - **S**creening **C**ompounds **D**irectory – 10,852,222 unique compounds
  - MDDR – 239,064 registered structures with bioactivity data
  - Toxicity – 172,542 registered structures
  - **C**omprehensive **M**edicinal **C**hemistry – 9603 registered structures
  - Metabolite – 71,359 molecules within 119,425 reactions

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

ACD

	Index	Name	Visible	Tagged	Visibility Locked	MDLNUMBER	CAS_TEXT	supplier.coden	CLOGP
1	1	Molecule	☑ Yes	☐ No	☐ No	MFCD00000001	645-96-5	ABCR ACC	2.29
2	2	Molecule	☐ No	☐ No	☐ No	MFCD00000002	7188-38-7	ABCR	
3	3	Molecule	☐ No	☐ No	☐ No	MFCD00000003	14542-93-9	ABCR ACC	
4	4	Molecule	☐ No	☐ No	☐ No	MFCD00000004			

ABCR
ACC

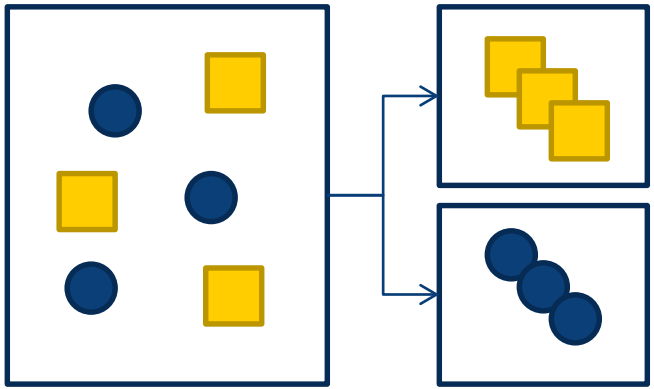
Total Products	Unique Compounds (3D Models†)	New Chemicals	Removed Chemicals*	Storage Requirements (ACD)			
				RCG Format		Direct Format **Top: Direct 8.0, Bottom: Direct 9.x**	
				Oracle Tablespace	Unzipped Dump Files	Oracle Tablespace	Unzipped Dump Files
28,539,359	**12,030,775 (11,988,413)**	**664,325**	**707**	**58.8 GB**	**38.0 GB**	**84.8 GB** **80.2 GB**	**51.5 GB** **48.7 GB**

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

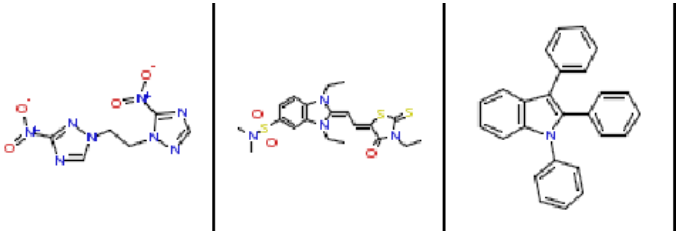
Copyright©2019 GGA Corp., All rights reserved.

Cluster subsets

- Number of molecule cluster
- Diversity of molecule cluster
- Countable property of molecule cluster
- ….



Save time for the sequential analysis steps

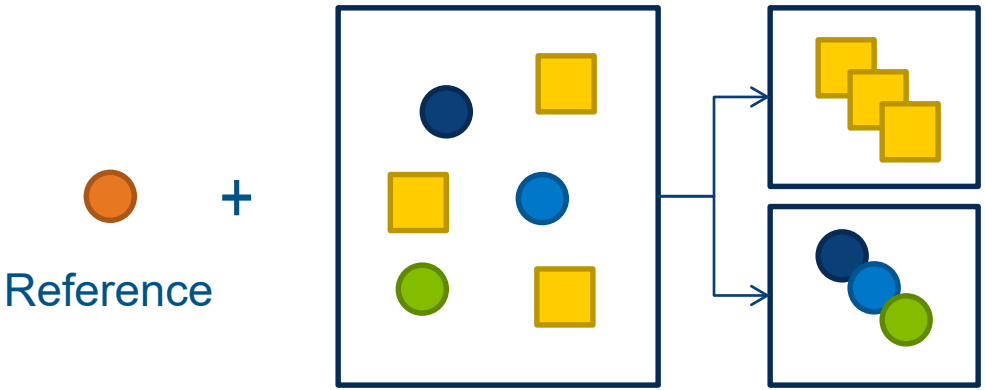


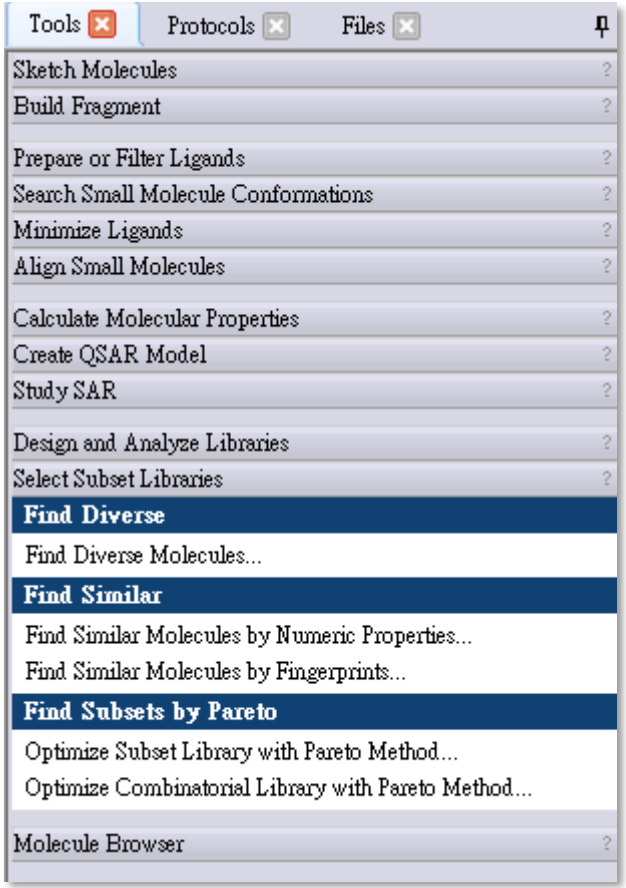
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技
GGA

Cluster subsets cont.

- Diversity of molecule cluster
- Countable property of molecule cluster



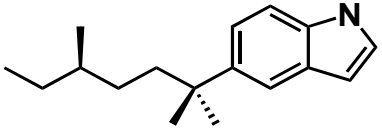


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技
GGA

Predict compound properties by Statistic method: QSAR

- Quantitative Structure-Activity Relationships



Chemical Data



IC_{50} 0.4 μM

Numerical Properties

Molecular Descriptor

MW	AlogP	Surface Area	Dipole	…
243.39	5.601	260.587	0.662	…

Define Correlation

$$IC_{50} = f([MW],[A\log P],\square)$$

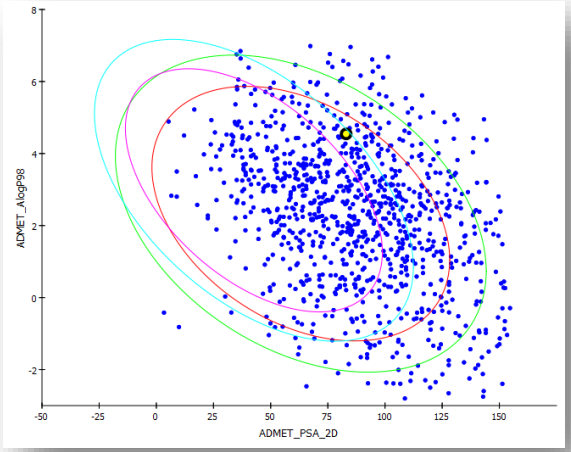
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

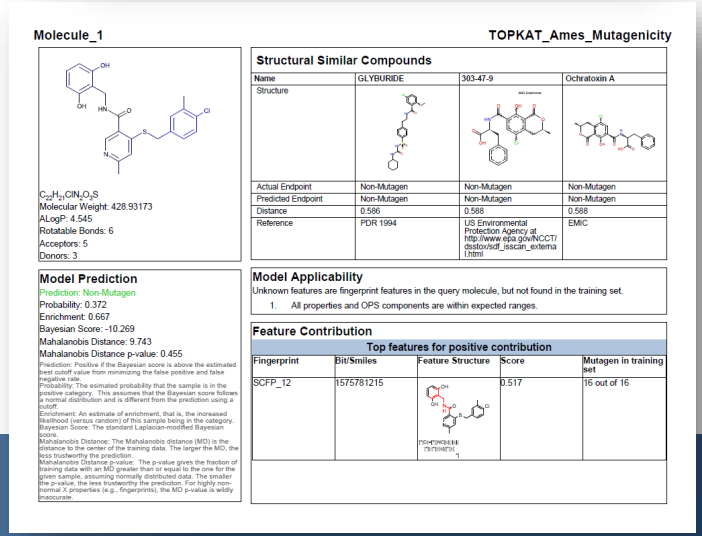
Copyright©2019 GGA Corp., All rights reserved.

ADMET and Toxicology

- ADMET
  - Human intestinal absorption
  - Aqueous solubility
  - Blood brain barrier penetration
  - Plasma protein binding
  - CYP2D6 binding
  - Hepatotoxicity
  - Filter sets of small molecules for undesirable function groups based on published SMARTS rules
- Property calculation
  - 2D and 3D molecular properties
  - Semi-empirical and DFT

- Predictive Toxicology
  - Ames mutagenicity
  - Rodent carcinogenicity (NTP and FDA data)
  - Weight of evidence carcinogenicity
  - Carcinogenic potency TD50
  - Developmental toxicity potential
  - Rat oral LD50
  - Rat maximum tolerated dose
  - Rat inhalation toxicity LC50
  - Rat chronic LOAEL
  - Skin irritancy and sensitization
  - Eye irritancy
  - Aerobic biodegradability
  - Fathead minnow LC50
  - Daphnia magna EC50
  - Log P





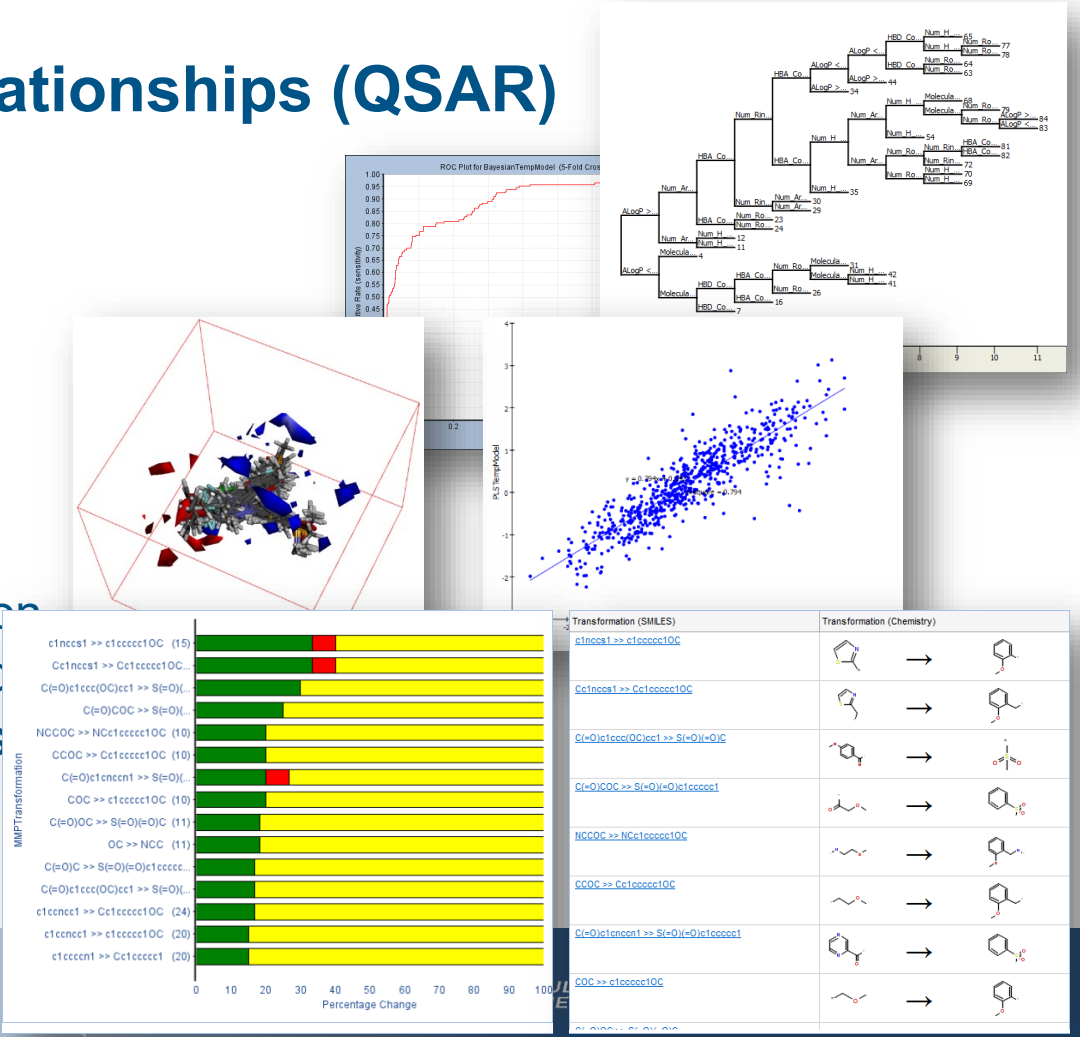
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

BIOVIA

創源生技 GGA

Structure-Activity Relationships (QSAR)

- Categorical data
  - Bayesian
  - Recursive Partitioning
- Continuous data
  - Genetic Function Approximation
  - Partial Least Squares
  - Multiple Linear Regression
- 3D molecular field-based
- Matched Molecular Pairs (MMPs) transformations and activity cliffs



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

BIOVIA

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Interface and Architecture

Copyright©2019 GGA Corp., All rights reserved.

Discovery Studio Product Architecture

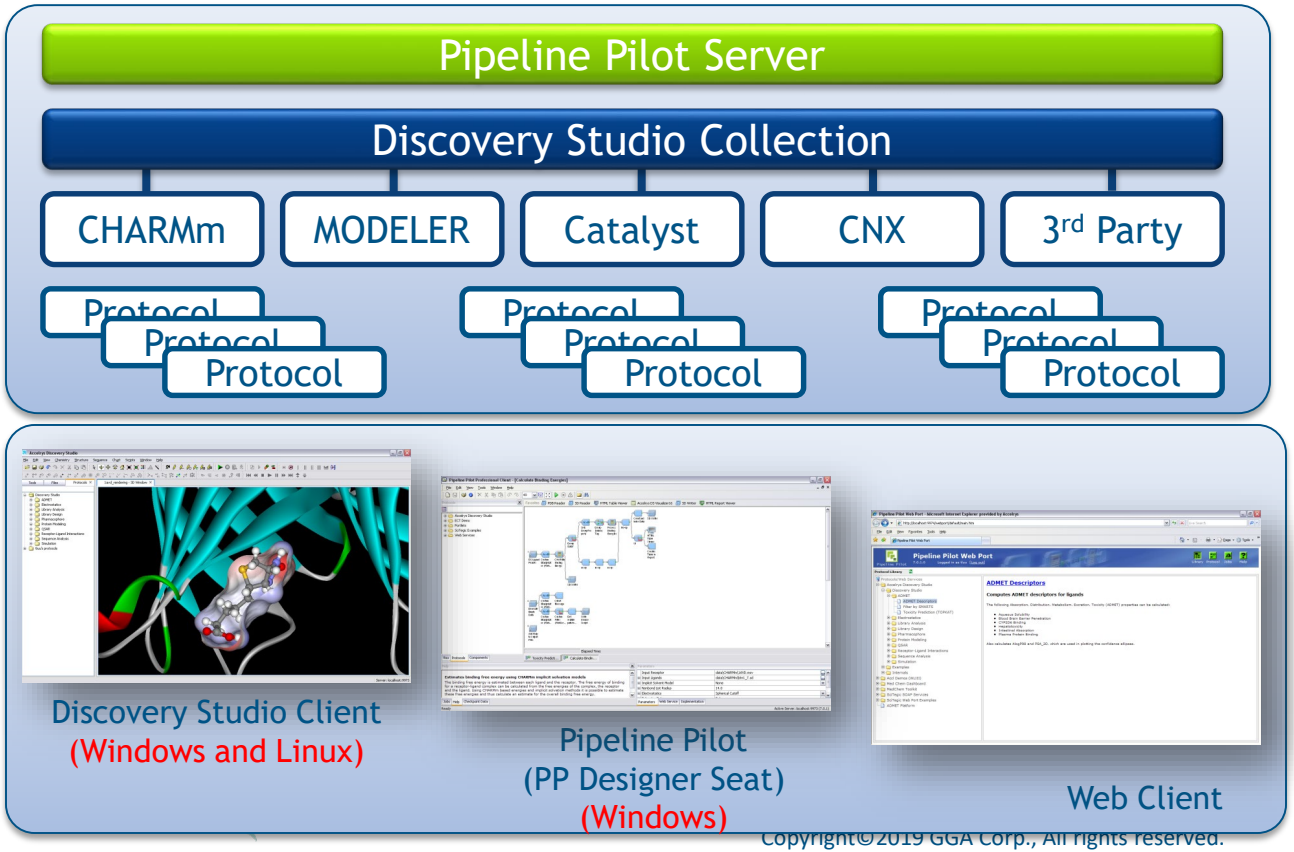
- Discovery Studio science runs on Pipeline Pilot:

- Server
  - AEP
  - PP Client
  - AEP Collection
  - DS collection
  - DS engines





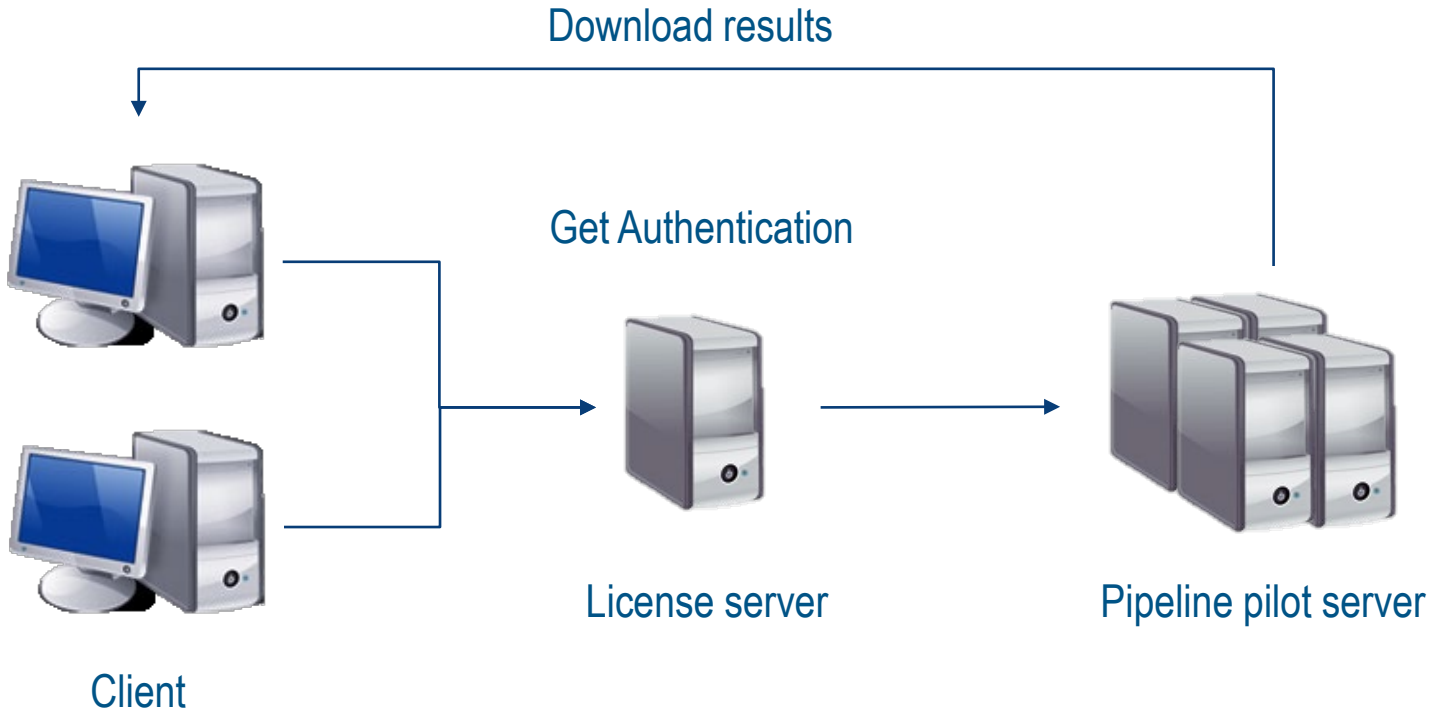
- Client(s)
  - DS client
  - PP Designer Seat
  - Web client



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright© 2019 GGA Corp., All rights reserved.

Discovery Studio Product Architecture

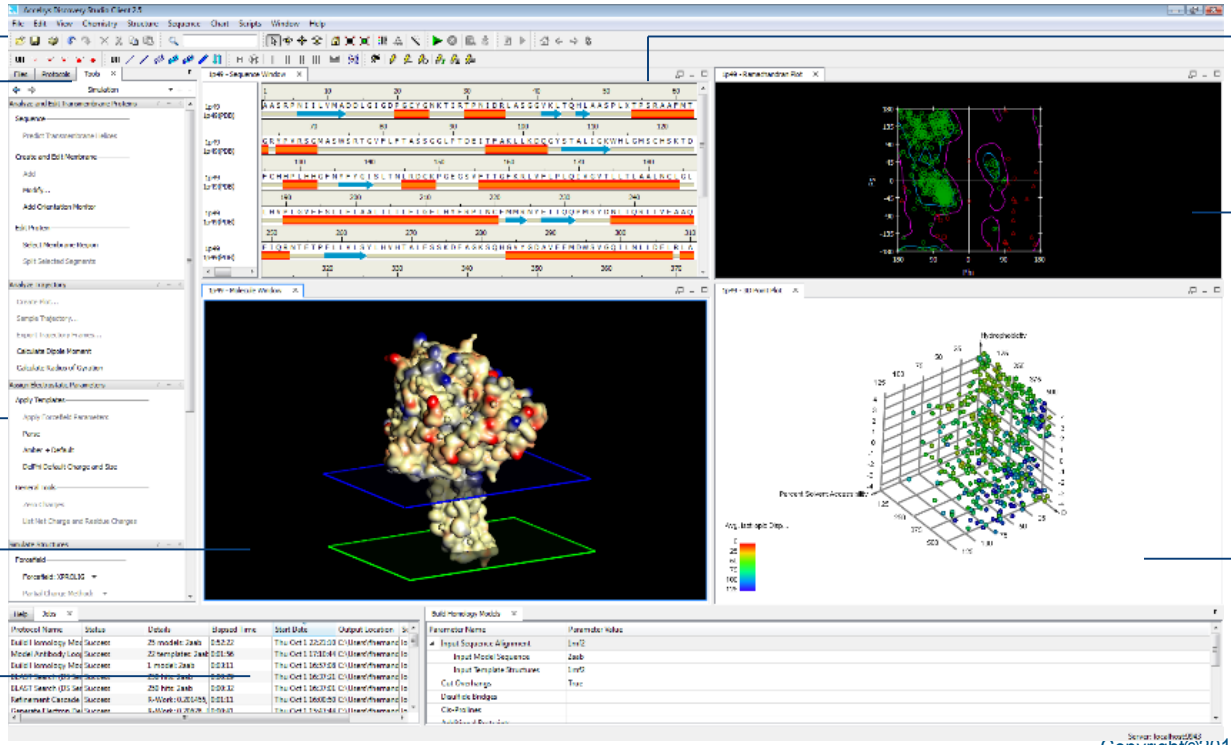


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技
GGA

Copyright©2019 GGA Corp., All rights reserved.

Supporting technology: Discovery Studio Client

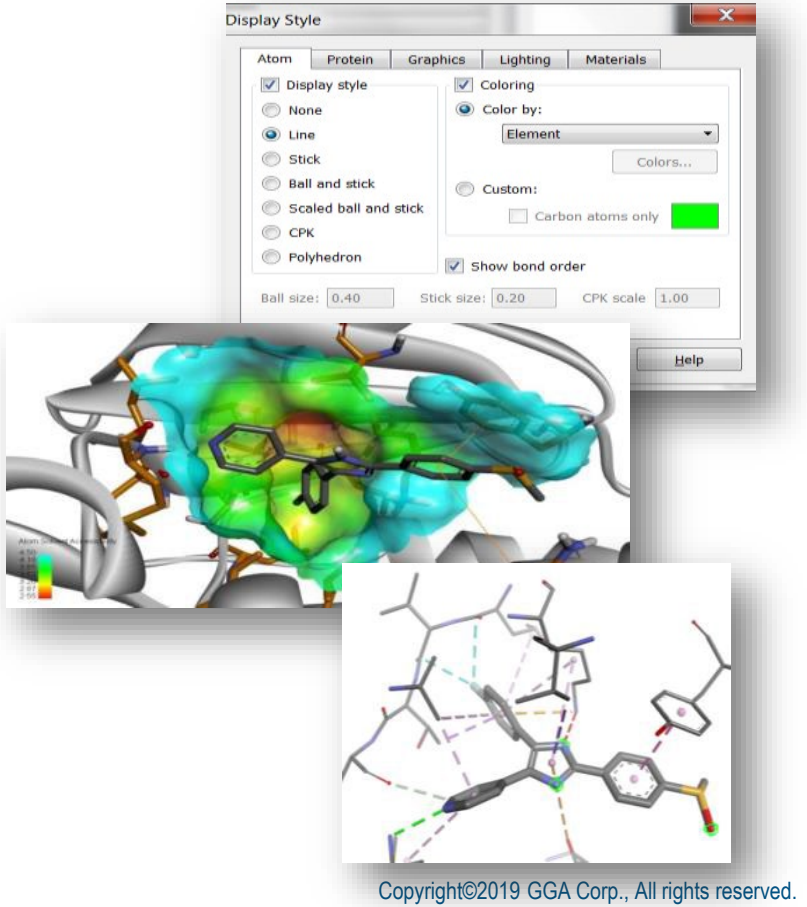


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Discovery Studio Client

- General
  - Display Style
    - Ball and Stick, CPK, Solid Ribbon
    - Surface/Protein Surface
    - Shadow
    - Clipping Plane
  - Monitors
    - Distance, Angle, and Bump
    - Interactions
    - RMSD

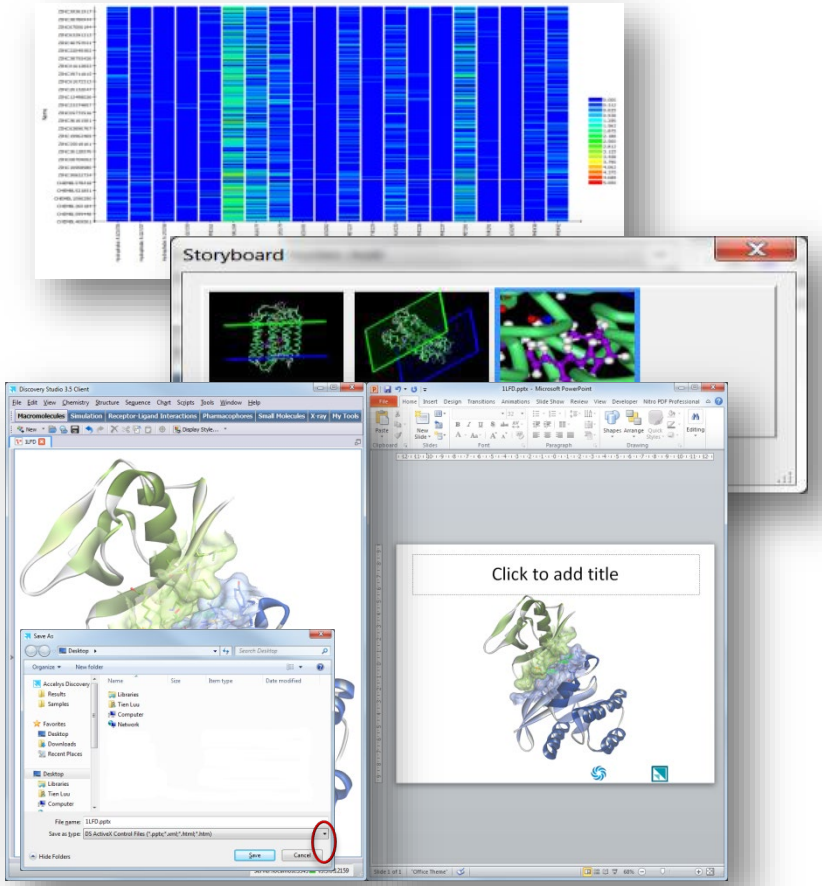


Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Discovery Studio Client

- General
  - Sequence view
  - Secondary structure prediction
  - Chart functions and data plots
    - Line plots, point plot, heat maps. etc
  - Collaboration, presentation functions
    - Story board, Active X
- Protocols
  - Automation and customization
  - Perl scripting

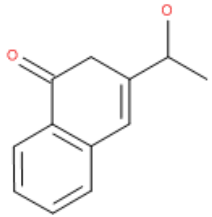


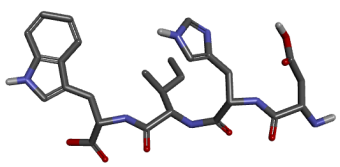
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

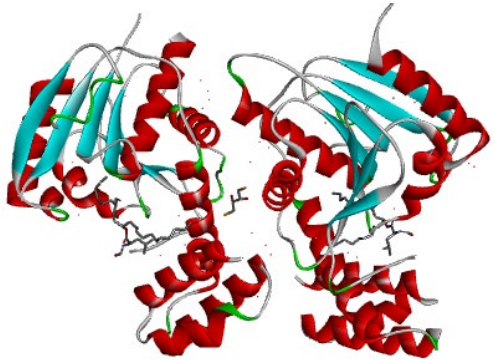
Data integration

- Open exist file
  - Chemical structure format
  - Sequence format
  - Protein structure format
- Download from database
  - Protein Data Bank (PDB)
  - NCBI Entrez Sequence Search
- Create new one
  - Small molecule, DNA, RNA, Peptide





```
>sp|P05067|A4_HUMAN Amyloid beta A4 protein OS=Homo sapiens GN=APP PE=1 SV=3
MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCGRLNMHMNVQNGKWDSDPSGTK
TCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG
EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR
GVEFVCCPLAEESDNVDSADAEEDDSDVWWGGADTDYADGSEDKVVEVAEEEEVAEVEEE
EADDDEDDEDGDEVEEEAEEPYEEATERTTSIATTTTTTTESVEEVVREVCSEQAETGPC
RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMAVCGSAMSQSLLKTTQEPLARD
PVKLPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEEAERQA
KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRRLALENYITAL
```

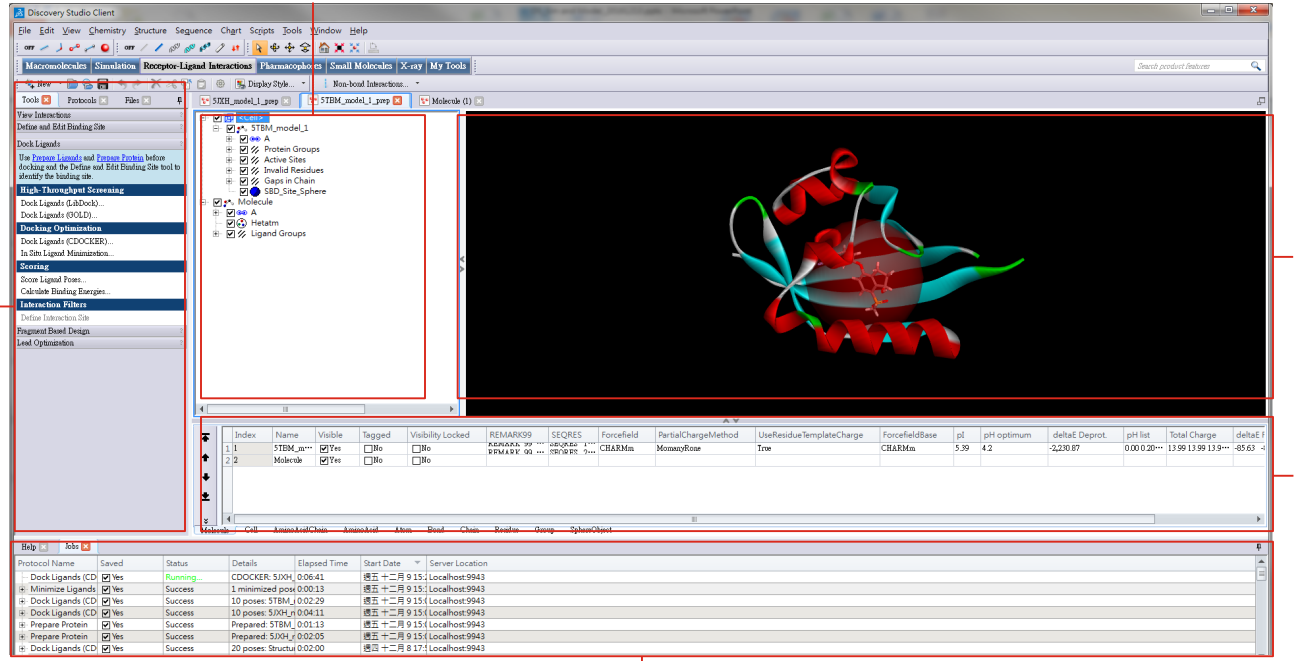


Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

DS Client - Windows

Hierarchy window



Tools
Protocols
Files

Molecule window

Table window

Job window

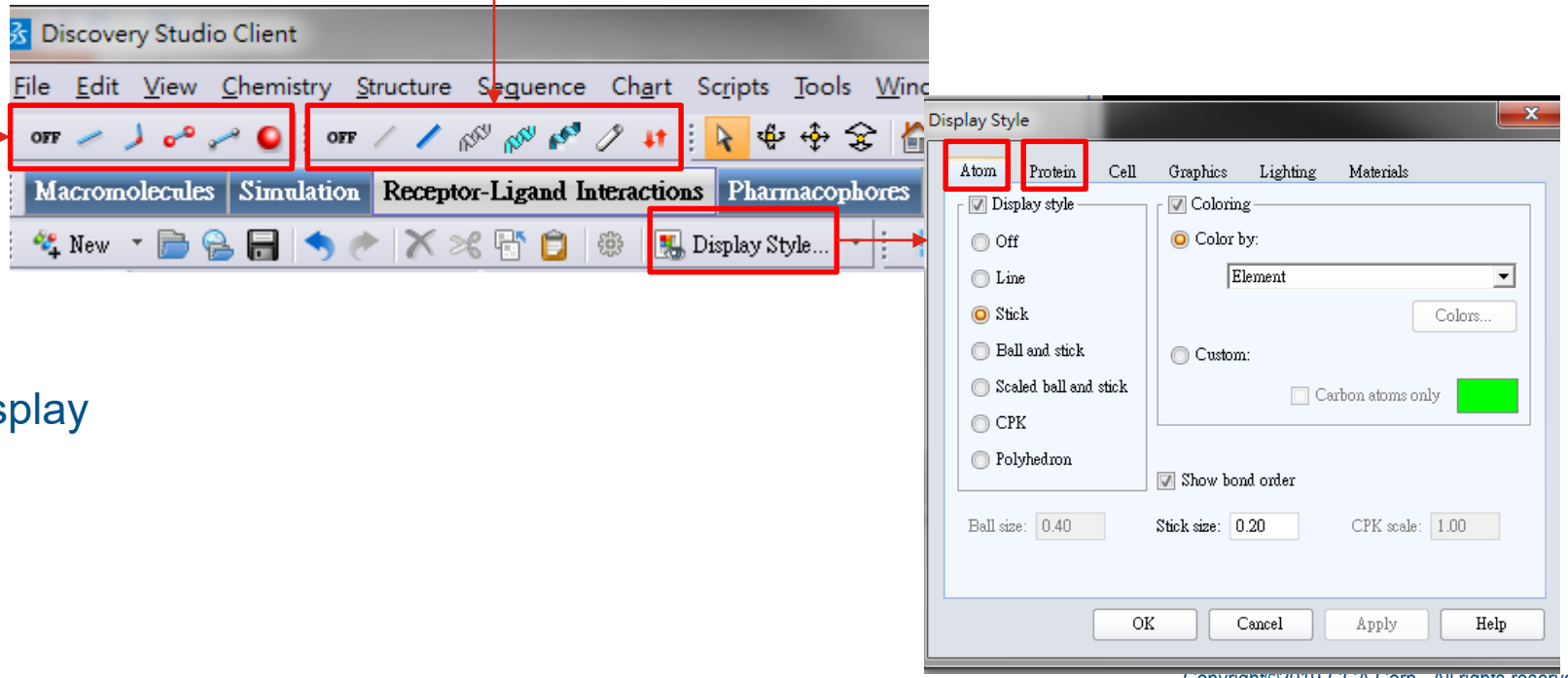
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

BIOVIA 創源生技 GGA DASSAULT SYSTEMES | The 3DEXPERIENCE® Company

Display style

Protein display

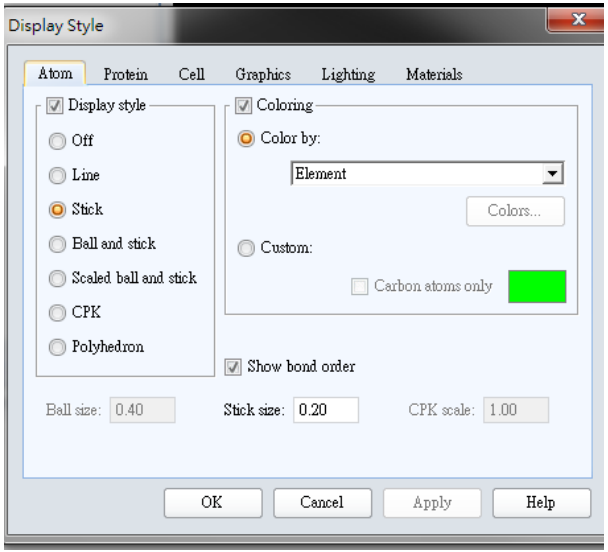


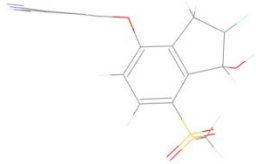
Atom display

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

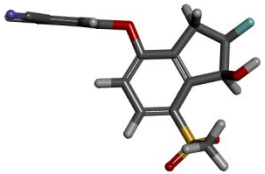
Copyright©2019 GGA Corp., All rights reserved.

Atom display

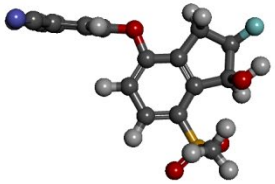




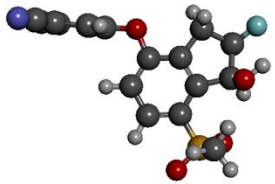
Line



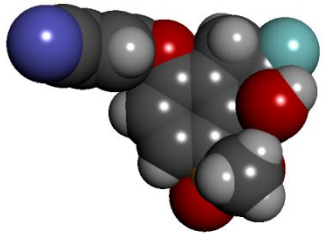
Stick



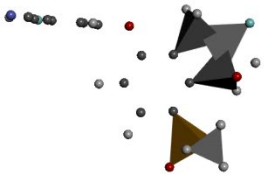
Ball and Stick



Scaled ball and stick



CPK

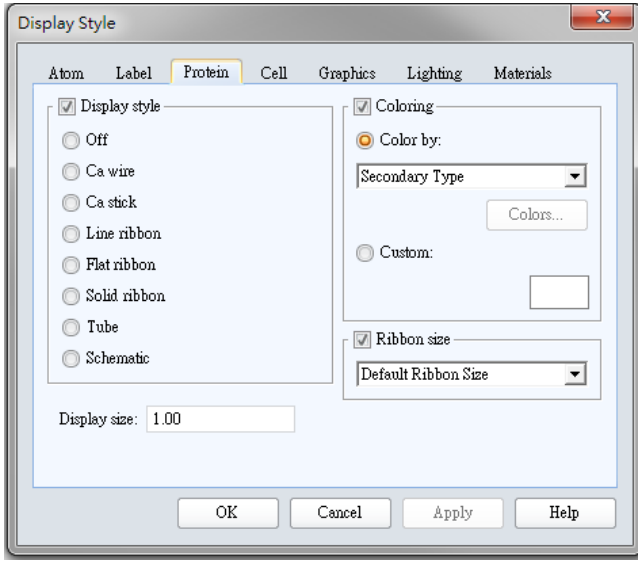


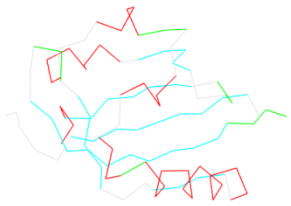
Polyhedron

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

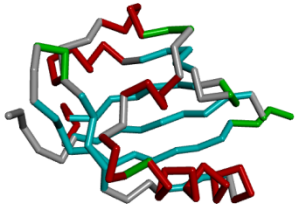
Copyright©2019 GGA Corp., All rights reserved.

Protein display

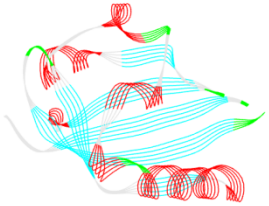




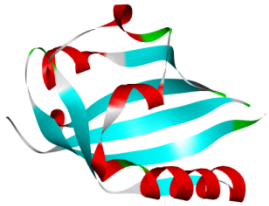
Ca wire



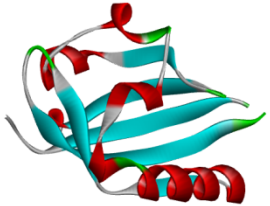
Ca stick



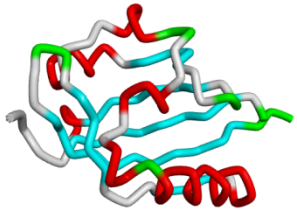
Line ribbon



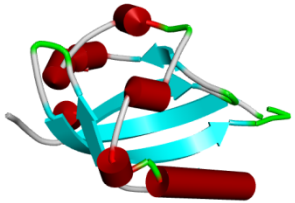
Flat ribbon



Solid ribbon



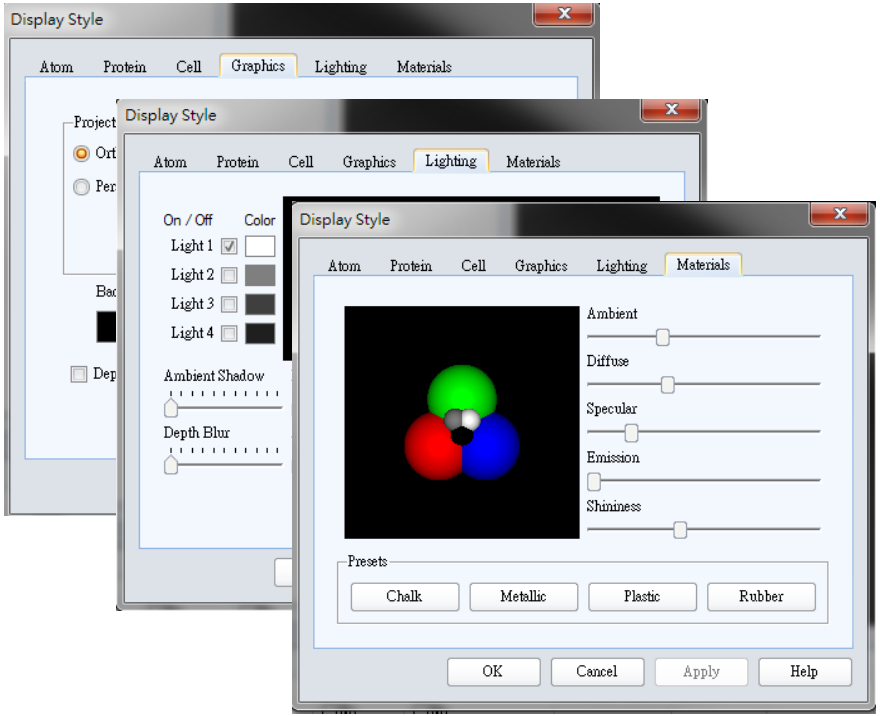
Tube

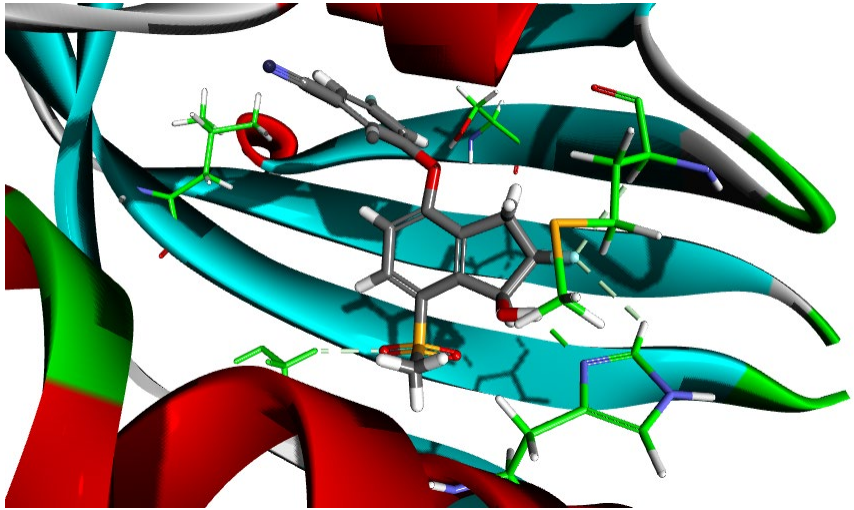


Schematic

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Display style cont.

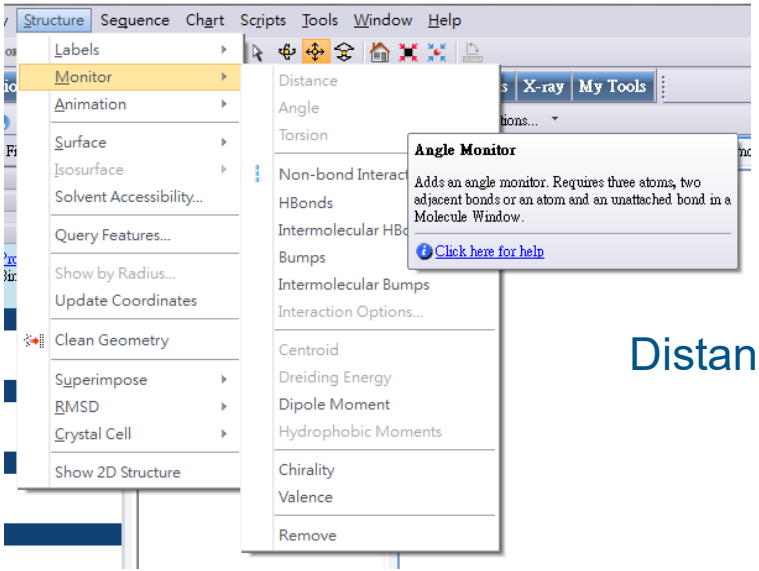




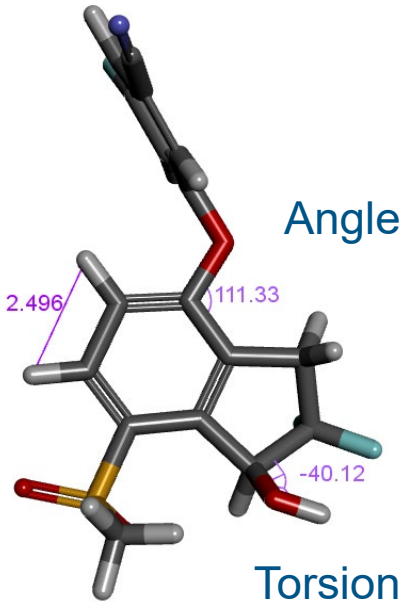
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Monitor and Non-bond Interactions tools



Distance



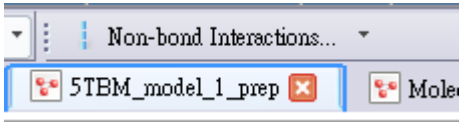
Angle

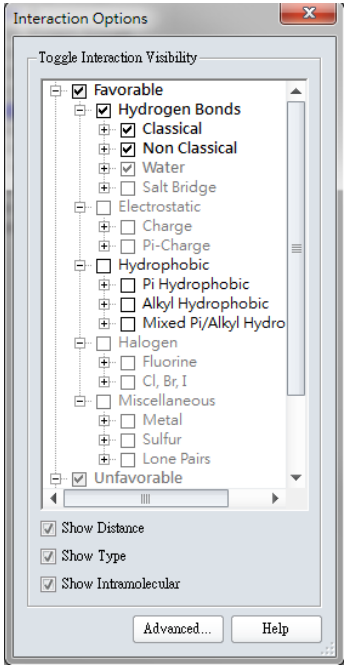
Torsion

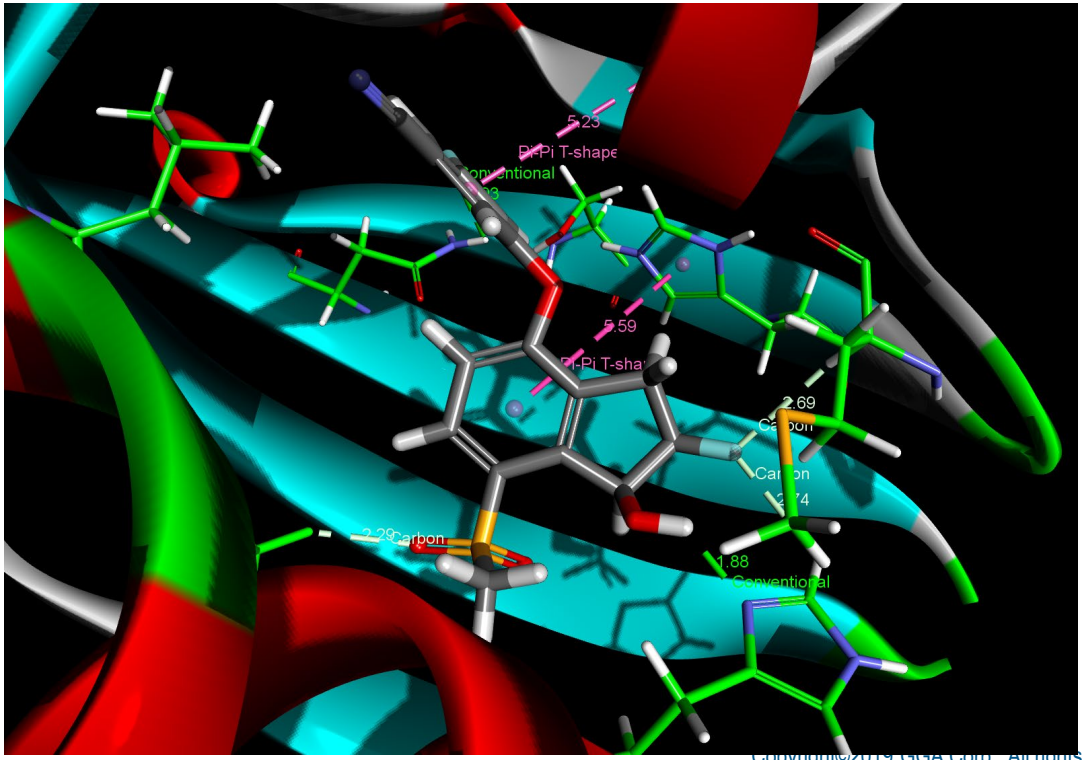
Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Monitor and Non-bond Interactions tools



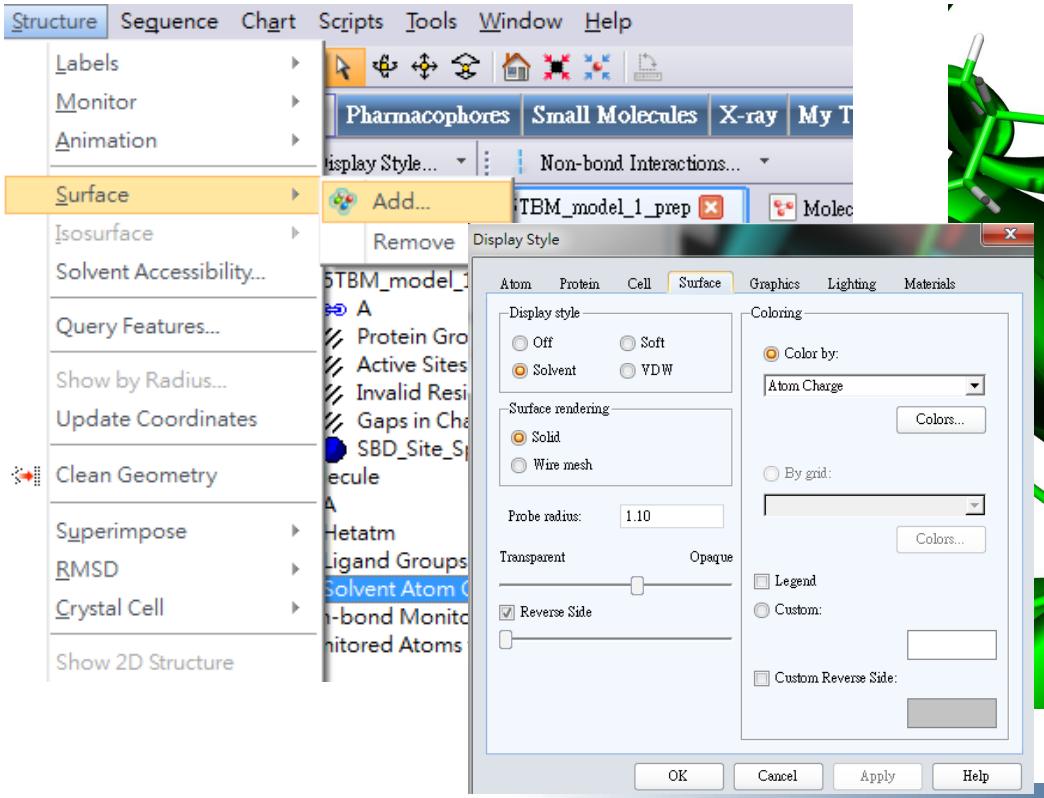


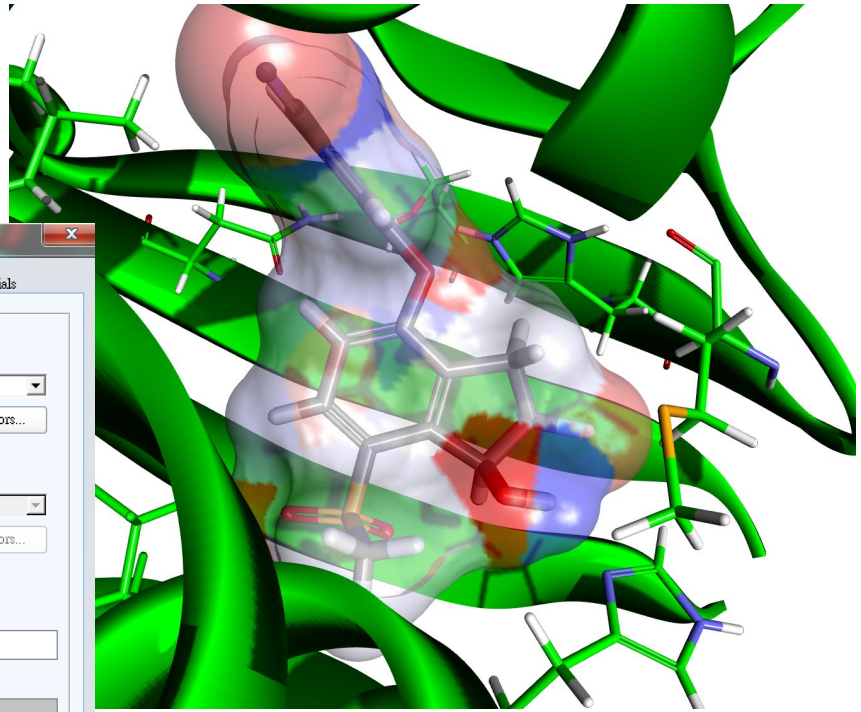


Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Surface

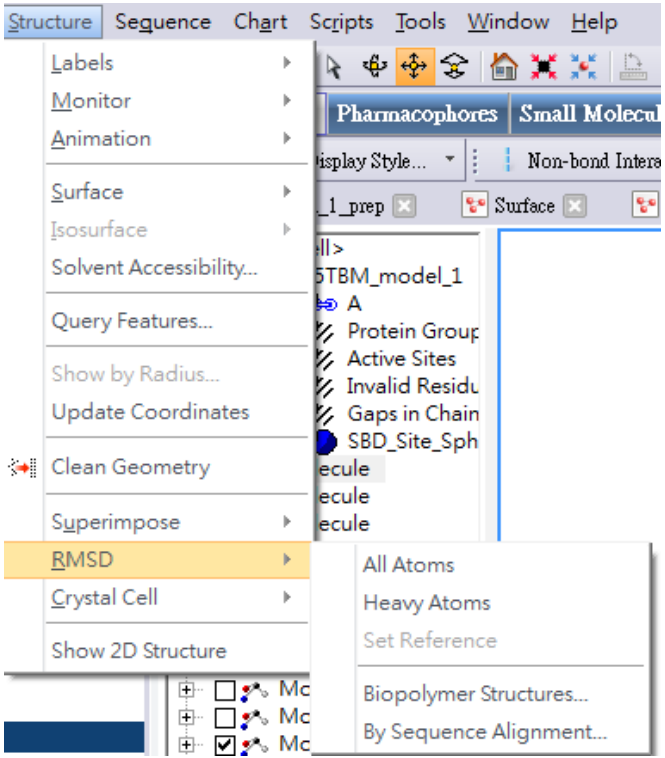


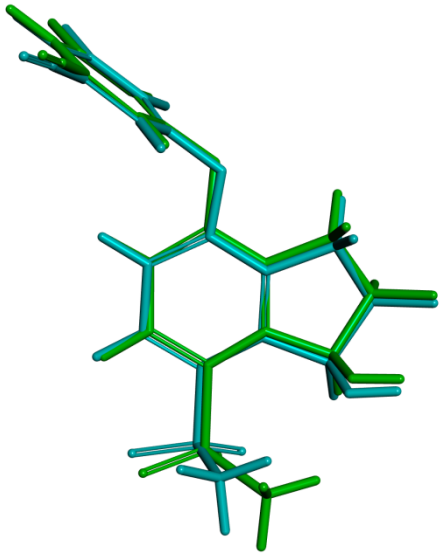


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 For MDU and NYHU

Copyright©2019 GGA Corp., All rights reserved.

Overlay/Superimpose/RMSD





Heavy Atom RMSD to Molecule 2

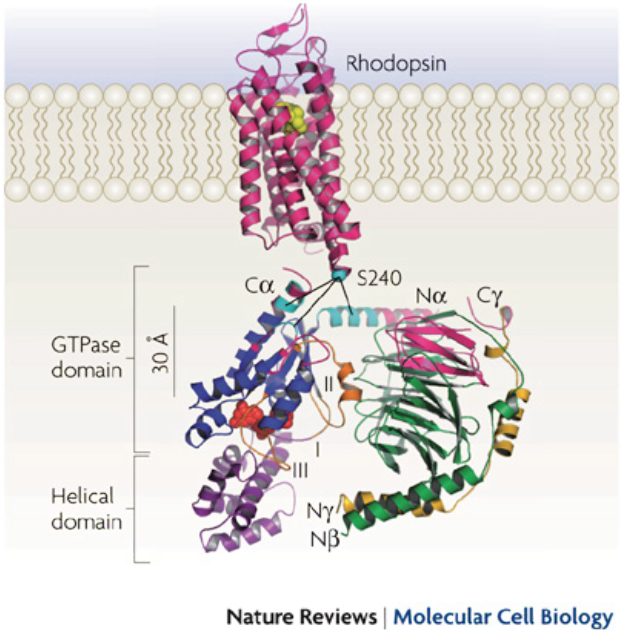
Name	Reference	RMSD (A)
Molecule 1	Molecule 2	0.4219

rved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Hands-on

- Try to change display style as the paper!
  - PBD code: 6n4b
  - A chain: Guanine nucleotide-binding protein G(i) subunit alpha-1
  - B chain: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1
  - C chain: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2
  - S chain: scFv16
  - R chain: Cannabinoid receptor 1



Nature Reviews | Molecular Cell Biology

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Agenda

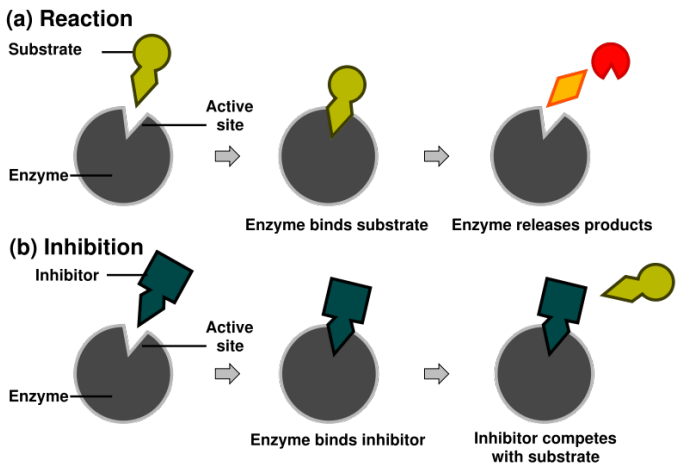
- **Introduction to Predictive Science**
  - Challenges of the Life Science Industry
  - Small Molecules and Macromolecule Solution
  - Hands-on 基本介面操作教學 (實機演練)
- **Discovery Studio: Receptor-Ligand Interactions**

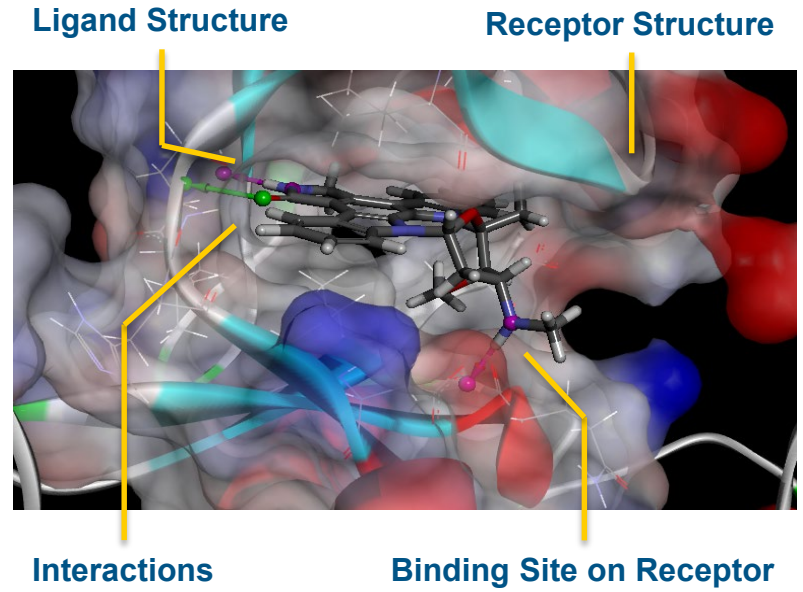
  小分子化合物接合模擬
  - Hands-on操作教學 (實機演練)

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking & Scoring – Structure-Based Design

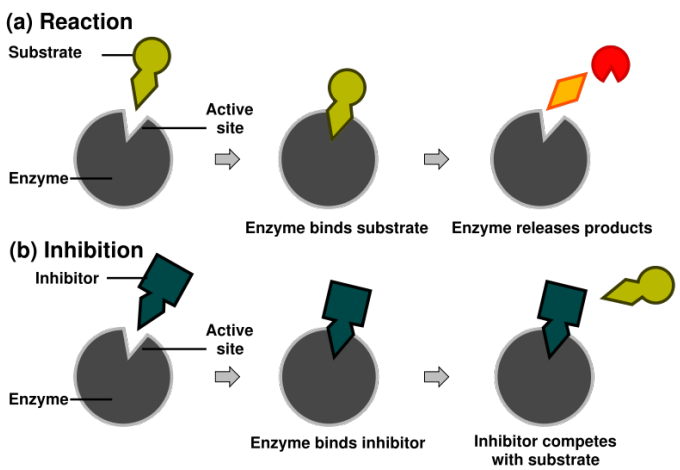


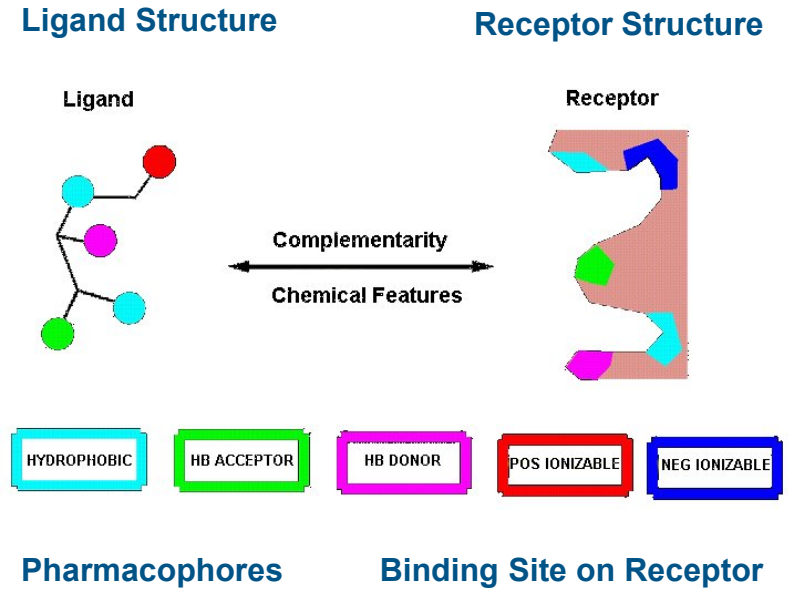


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking & Scoring – Pharmacophore-Based Design

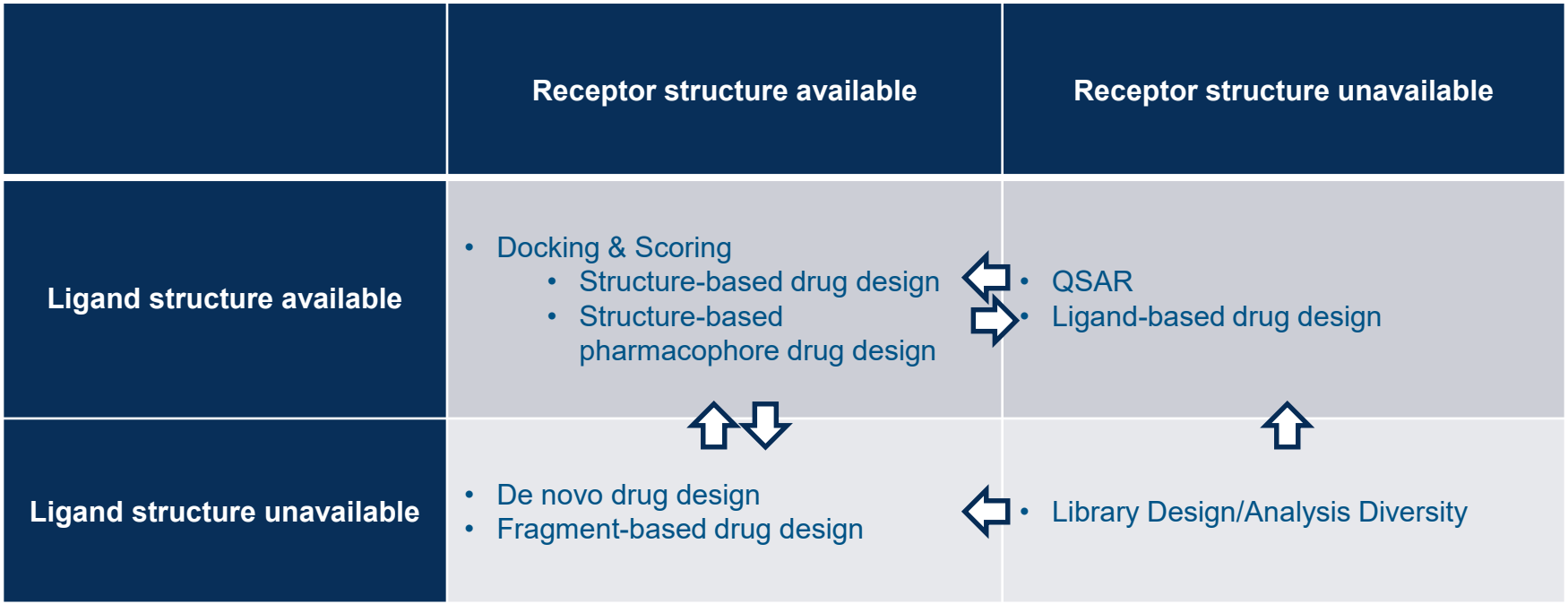




3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Strategy for Small Molecule Drug Design



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Strategy for Small Molecule Drug Design

	Receptor structure available	Receptor structure unavailable

Ligand structure available	• **Docking & Scoring**
	  • **Structure-based drug design**
	  • **Structure-based pharmacophore drug design**
		• QSAR
		• Ligand-based drug design

Ligand structure unavailable	• De novo drug design • Fragment-based drug design	• Library Design/Analysis Diversity

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Structure-Based Drug Design

- Using knowledge of a receptor to guide design of new ligands
  - Structure of the receptor is known
  - Can use either an experimental or homology model

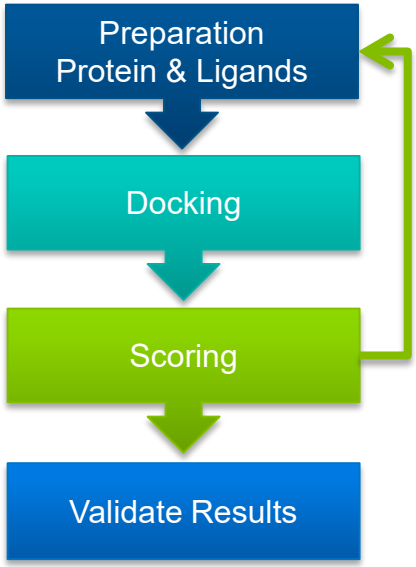
- One approach is to identify potential ligands that can bind to receptor
  - High-throughput virtual screening
  - Rigid or flexible docking
  - Scoring of docked ligands

- May need to identify the binding site
  - Active site search

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking Workflow



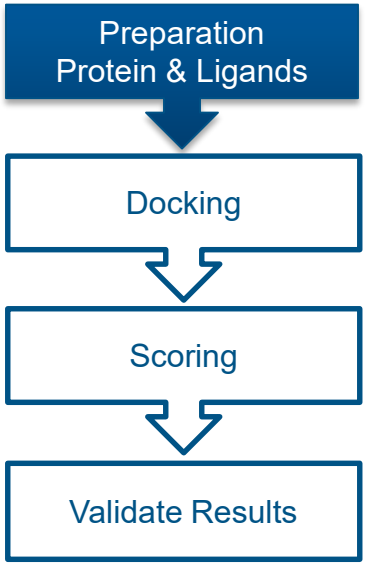
Repeat Cycle = Virtual High Throughput Screening (vHTS)

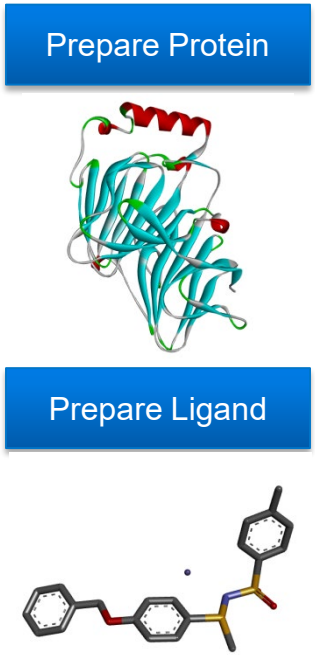
- In order to maximize the rate of success, the evaluation phase in each steps is very important.
- Requirements
  - Known active compounds and decoy
  - Bound molecules
  - Optimize the parameters of docking technique
  - Prioritized the molecules in library before proceeding the docking and scoring steps

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking Workflow





Protein Preparation

- Repair deficient structure
- Identify binding site
- Modify the protonation state

Ligand Preparation

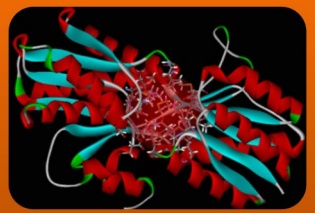
- Add hydrogen atom
- Generate 3D structure coordinates
- Creating isomers
- Remove duplicates
- Valence modification
- Standardization charges

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

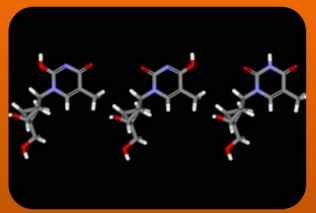
SBD: Input Preparation

Proteins



- Standardise atom names
- Insert missing atoms in residues
- Remove alternate conformations
- Insert missing loops
- Optimize short & medium size loops
- Calculate pK and protonate

Ligands



- Add hydrogens
- Calculate 3D coordinates
- Enumerate ionization states
- Ionize functional groups
- Generate tautomers and isomers
- Remove duplicates
- Fix bad valencies
- Standardize charges for common groups
- Retain largest fragment

Fragments



- Generate fragments using RECAP* rules
- Rule of Three filters

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Selecting the Protein Receptor

- Choosing the protein receptor
- Probably have a receptor in mind already
  - Based on biological or medical problem
  - Reinforced by biological data
- Requires a three-dimensional structure of the receptor
  - X-ray crystal structure
  - NMR structure
  - Homology model
- Can be an apo form of receptor

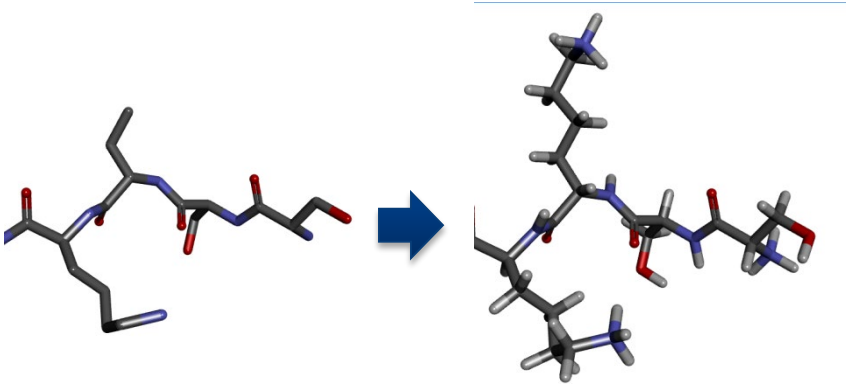
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Protein Preparation

- Before any docking can be performed, the receptor must be properly prepared
  - Particularly a concern with PDB files

- Preparation includes having...
  - All residues completed
  - Correct chemistry
    - Correct bond orders
    - Correct atom valences
  - All required hydrogen atoms added
  - Correct formal charges

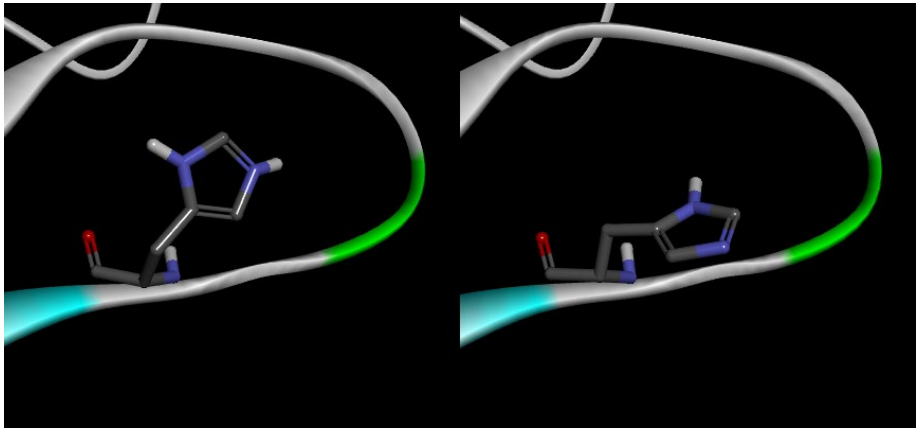


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Protonation state

- Calculate pKa for each residue in different pH
- Modify the protonation (ionization) status



Example for His334 of Protein A in different protonation status

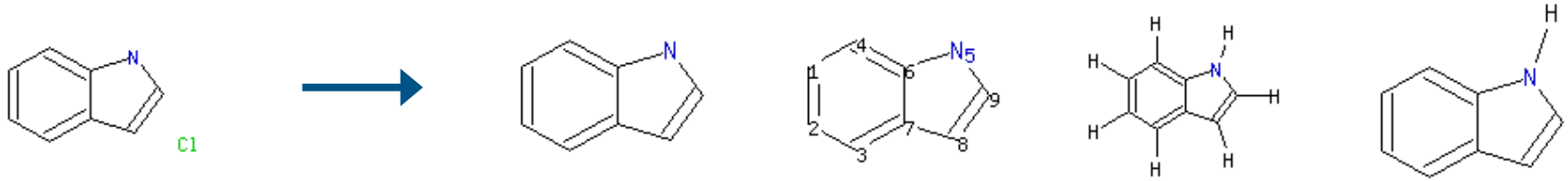
Protein A	pH7.2	pH8.0
HIS334	1	0
HIS376	1	0
HIS1253	1	0
HIS1540	1	0
HIS1575	1	0
HIS1591	1	0
HIS1937	1	0
HIS2391	1	0
HIS2452	1	0
HIS2455	1	0
HIS2488	1	0
HIS2500	1	0
HIS2625	1	0
HIS2826	1	0
HIS2865	1	0
HIS2898	1	0
HIS2946	1	0
HIS2998	1	0
GLU1018	0	-1

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Standardize molecules

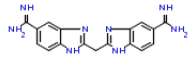
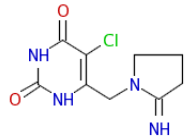
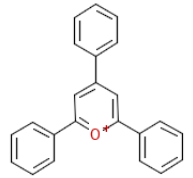
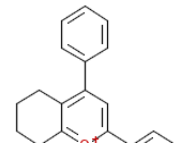
- Beautification
  - Keep/Remove largest/smallest fragments (option of Standardize Molecule)
  - Add/Remove atom numbers (found in Utilities)
  - Add/Remove hydrogens
  - Add hetero hydrogens (option of Add Hydrogens component)
  - Center molecule (option of Standardize Molecule)



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Standardize molecules cont. - Strip salts

- *Strip Salts* will strip a parent molecule of its counter ions.
- Chemistry Data\Queries\Salts.sd contains any defined salt structure
- User defined salt queries can be added via parameter User Salts
- Further components: *Identify Salts* and *Generate Salts*

Molecule	CODE	RemovedSalts
	SPB 08110	HCl HCl HCl HCl H2O
	SPB 08153	HCl H2O
	TL 00001	[O-]S(=O)(=O)C(F)(F)F
	TL 00002	[O-]S(=O)(=O)C(F)(F)F

Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Ligand preparation

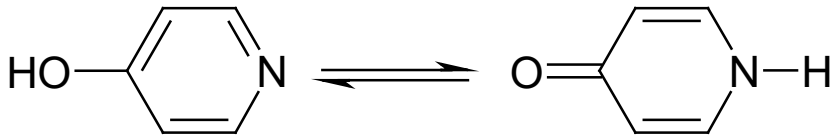
Performs the following steps, some of which can be controlled by the protocol parameters:

- Generate a canonical tautomer
- **Keep only the largest fragment**
- **Set standard formal charges on common functional groups**
- Kekulize the molecule
- **Enumerate ionization states at a given pH range (Optional)**
- Enumerate tautomers (Optional)
- Enumerate isomers (Optional)
- By default only unspecified bonds and atoms are enumerated
- **Remove duplicate structures (Optional)**
- Filter structures that violate Lipinski rules (Optional)
- Generate a standard 3D conformation (Optional)
- Catalyst is used to generate a reasonable 3D conformation

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Tautomers



- Compounds whose structures differ markedly in arrangement of atoms, but which exist in easy and rapid equilibrium, are called tautomers.

  This means:

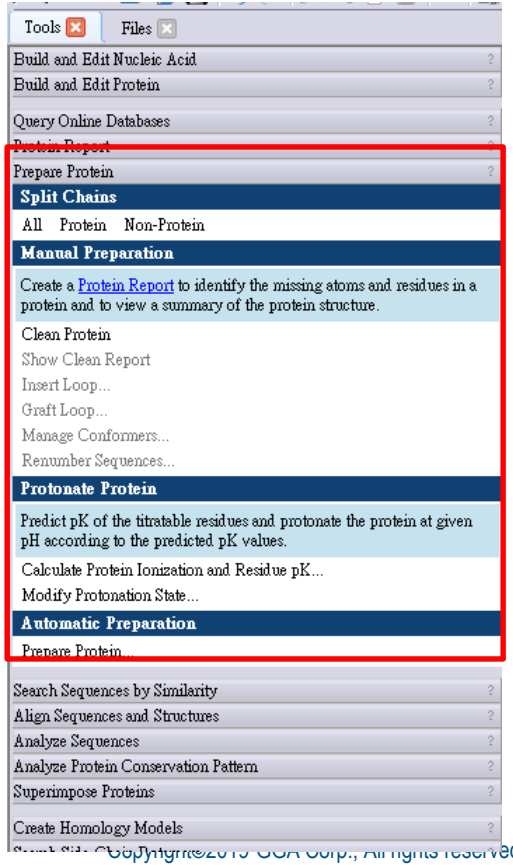
  – Possible duplicate molecules may be missed.
  – Different structures may have different values for calculated properties.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

創源生技
GGA

Protein Reports and Utilities Tools

- Allows you to:
  - Renumber sequences
  - Summarize information about the protein
  - Generate hydrophobicity plots
  - Split structures into distinct molecules
  - Clean protein molecules
  - Add missing atoms
  - Fix connectivity
  - Fix names
  - Define a template for a nonstandard amino acid

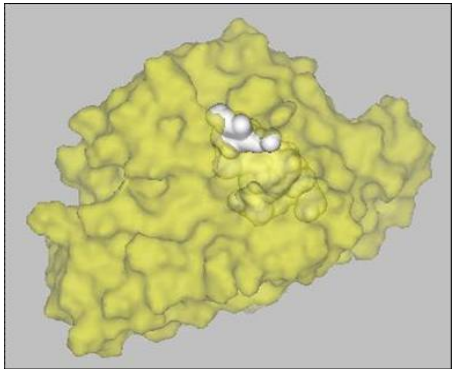


Copyright© 2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

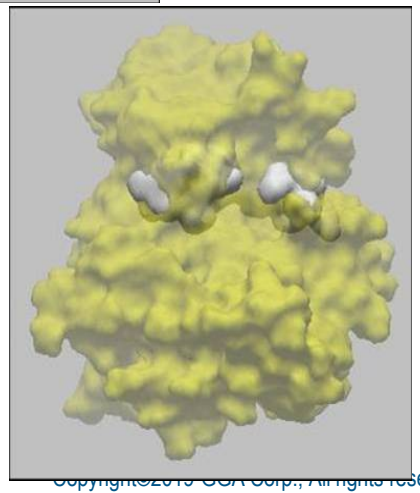
Binding Site Identification

- Liang et al. 1998 found small molecule binding sites to be:
  - Indentations, crevices, or cavities
  - And often the largest site is the true binding site
- Laskowski et al. 1996 reported an analysis of cleft volumes:
  - Often the ligand is bound in the largest cleft
  - Usually the largest cleft is considerably larger than the others
- Jones and Thornton 1997 found that protein-protein interaction sites tend to be:
  - Flat and hydrophobic



HSV-1 thymidine kinase

Abl tyrosine kinase



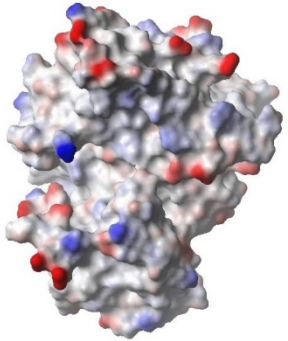
Copyright© 2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Binding Site Identification

- With experimental structure determination of ligand complexes...
  - Binding site is often identified
  - Analogues can show additional features
  - Use the position of known ligands to limit possible binding sites for new candidates

- With unbound proteins…
  - Binding site may not be obvious
  - Binding site must be sought





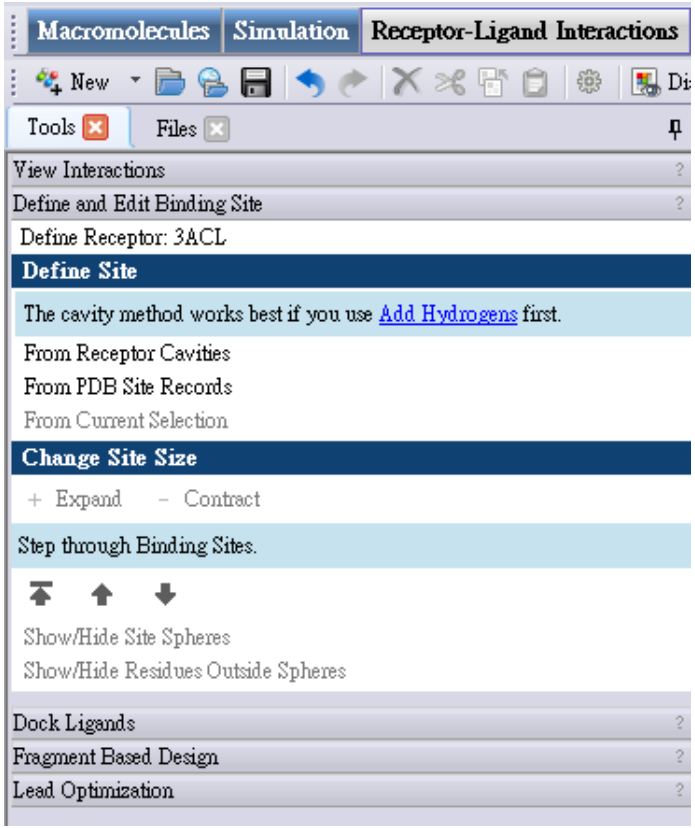
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Binding Site Identification

- When the binding site is unknown…
  - Search for cavities
  - Use experimental data
    - Site-directed mutagenesis studies
    - Cross-linking data
    - NMR results
  - Compare target to similar proteins

- Can be accomplished with Binding Site Tool panel

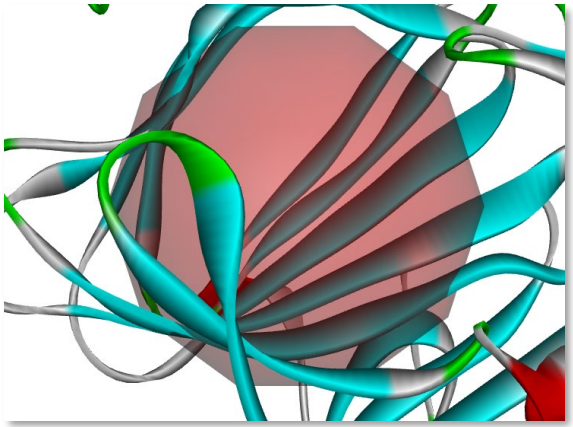


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

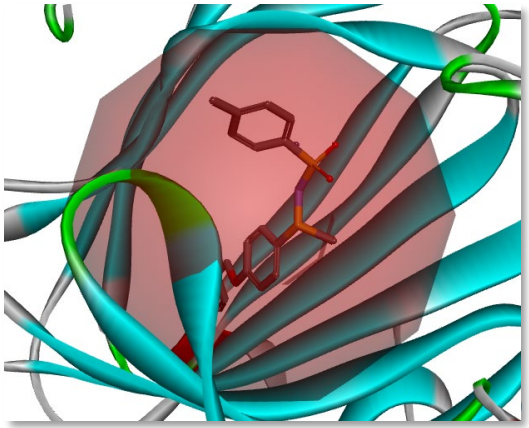
Copyright©2019 GGA Corp., All rights reserved.

Site Search Approaches

- Protein Shape
  - Based on the shape of the protein only
  - Identifies cavities and crevices



- Bound Ligand Volume
  - Requires presence of a bound ligand in protein
  - Identifies region around bound ligand

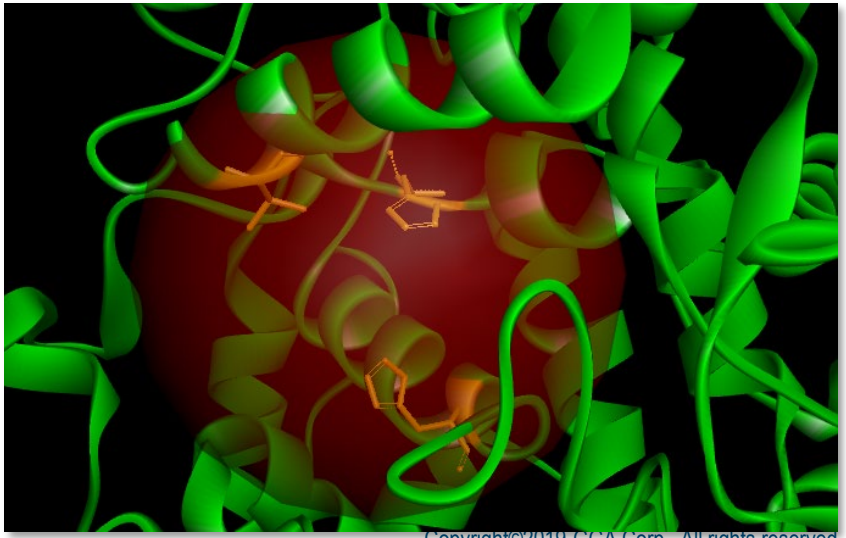


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

GGA Corp., All rights reserved.

Site Search Approaches

- Use experimental data
  - Identified binding site from references
  - Site-directed mutagenesis studies
  - Compare target to similar proteins

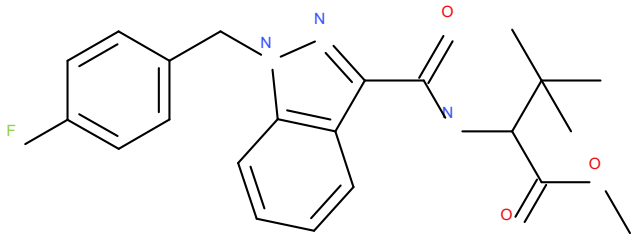


Copyright©2019 GGA Corp., All rights reserved.

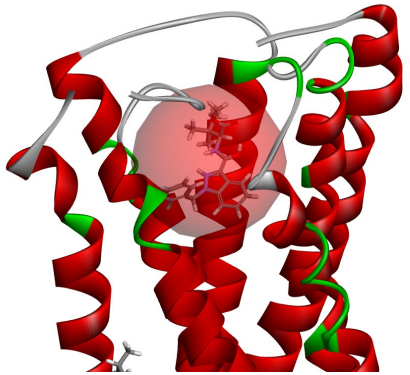
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Demo

- Prepare Protein – 6N4B
- Prepare Ligand - MDMB-Fubinaca
- Define Binding Site
- Remove in-situ Ligand (MDMB-Fubinaca)
- Dock computational MDMB-Fubinaca structure to binding site



MDMB-Fubinaca

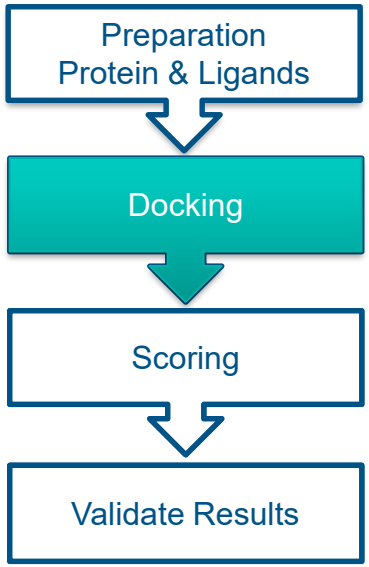


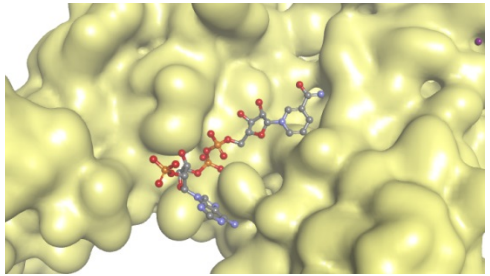
PDB Code: 6N4B

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking Workflow





Docking

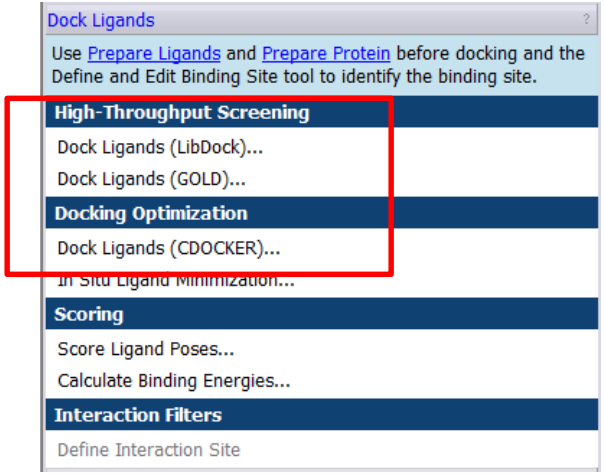
- Generate conformations
- Find a pose
- Docking tools in DS
  - Dock Ligands (CDOCKER)
  - Dock Ligands (LibDock)
  - Dock Ligands (GOLD, need extra license)
  - Dock Ligands (LigandFit)
  - Pharmacophore Docking
  - Flexible Docking

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

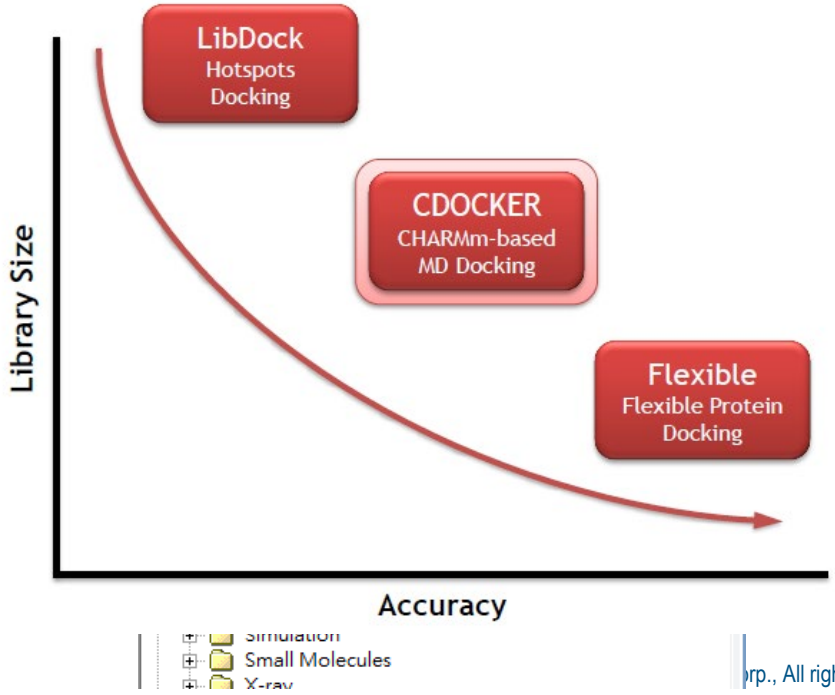
Copyright©2019 GGA Corp., All rights reserved.

Docking from Tool Panel

- Docking tools in DS
  - Dock Ligands (CDOCKER)
  - Dock Ligands (LibDock)
  - Dock Ligands (GOLD)



- Docking tools in DS (Protocol)



3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

rp., All rights reserved.

Comparative performance of LibDock and CDOCKER on AstexDiverse dataset

Method	% Docked Accurately (Best)	% Docked Accurately(Top)	RMSD Average (Best)	RMSD Average (Top)	Tim(min)
LibDock	91	50	1.2	3.8	0.5
CDOCER	94	79	0.8	1.5	5.0

LibDock is optimized for speed:
get accurate docked poses in seconds

**CDOCKER is optimized for accuracy:
get significant improvement in rank-ordering of correct pose and RMSD to X-ray structure**

1. Hartshorn, et al. J. Med. Chem., 50 (4), 726 -741 (2007)
3 GHz

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

CDOCKER



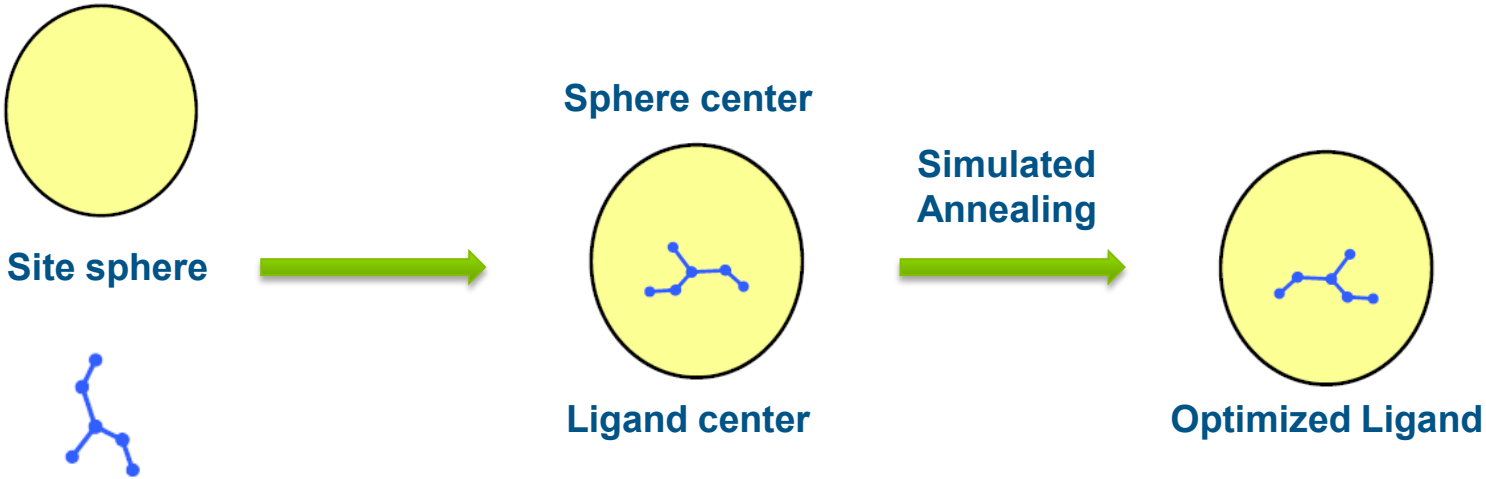
- CDOCK is a grid-based molecules docking method.
- Ligand conformations are obtained by Molecular Dynamic (MD) methods

Wu G, Robertson DH, Brooks CL III, Vieth M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER - A CHARMm-based MD docking algorithm. *J. Comp. Chem.* **2003**, 13, 1549

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

CDOCKER : Ligand Fitting



Ligand

The conformation generated by high temperature MD

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Small Molecule Docking using CDOCKER

- 41 protein-ligand complexes from the PDB
  - Structurally diverse set of ligands
  - All-atom representation used in published work[a]
- Grid-based approach used in Discovery Studio
  - Faster method
  - No significant compromise on accuracy

a. Erickson et al. J Med Chem (2004) 47:45-55

Success rates and CPU times for each algorithm

Docking Algorithm	success rate	av/median CPU times
GOLD[b]	46.3	824/708[c]
Dock[b]	51.2	200/114[c]
FlexX[b]	53.7	65/35[c]
CDOCKER (No Grid)[b]	82.9	2012/1740[c]
Discovery Studio CDOCKER (Grid)	78.1	630/580[d]

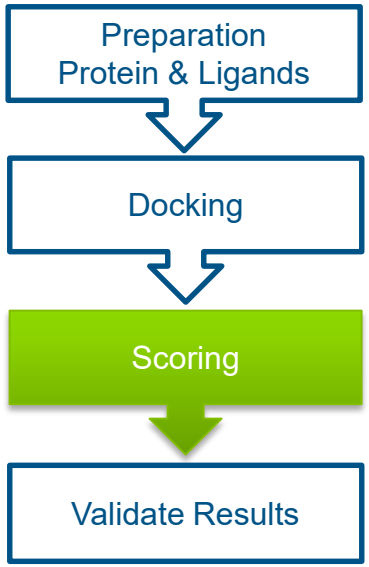
Single best docking run for each algorithm
(Success is defined as RMSD < 2Å from X-ray structure)

# of rot. bonds	total # of ligands	# of ligands docked correctly				
		Dock[b]	GOLD[b]	FlexX[b]	CDOCKER (No Grid)[b]	Discovery Studio CDOCKER (Grid)
< 8	20	18	14	16	19	18
≥ 8	21	3	4	6	15	14

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking Workflow



Literature Scoring Functions

- Also known as Empirical/Knowledge-based scoring functions
- based on counting the number of various types of interactions between the two binding partners
- based on statistical observations of intermolecular close contacts in large 3D databases
- Scoring functions
  - LigScore 1 & 2
  - Piecewise Linear Potential (PLP) 1 & 2
  - Potential of Mean Force (PMF) & PMF04
  - Jain
  - Ludi 1, 2 , & 3

Energy-Based Functions

- Binding energy calculation method based on the following formula

$$E_{binding} = E_{complex} - (E_{ligand} + E_{receptor})$$

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

erved.

Scoring

- Scoring of docked poses is still a major challenge
- Aim of scoring:
  - Identification of the correct binding pose by lowest energy value
  - Ranking of protein-ligand complexes according to their binding affinities
- No single scoring function can correctly rank every protein-ligand complex
  - Relative contribution of different protein-ligand interactions may vary between structural families
- Use consensus scoring
  - Combination of several scoring functions

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Types of Scoring Functions

- Empirical scoring functions
  - Derived from training sets of protein-ligand complexes with determined affinity data
- Force field-derived functions
  - Handle the ligand binding prediction with the use of potential energies (non-bonded interaction terms)
  - Could include solvation and entropy contributions
- Knowledge-based functions
  - Based on atom pair potentials derived from structural databases
  - Forces and potentials are collected from known protein-ligand complexes to get a score for their binding affinities

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

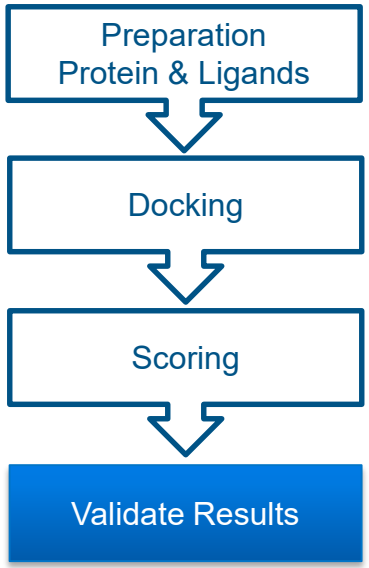
Scoring Ligand Poses Protocol

- Used for final evaluation of docked poses
- Varity of scoring functions available
  - LigScore1/LigScore2
  - PLP1/PLP2
  - Jain
  - PMF/PMF04
  - Ludi
- Can be combined in Consensus Scoring protocol
- Can also be applied during docking

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Docking Workflow



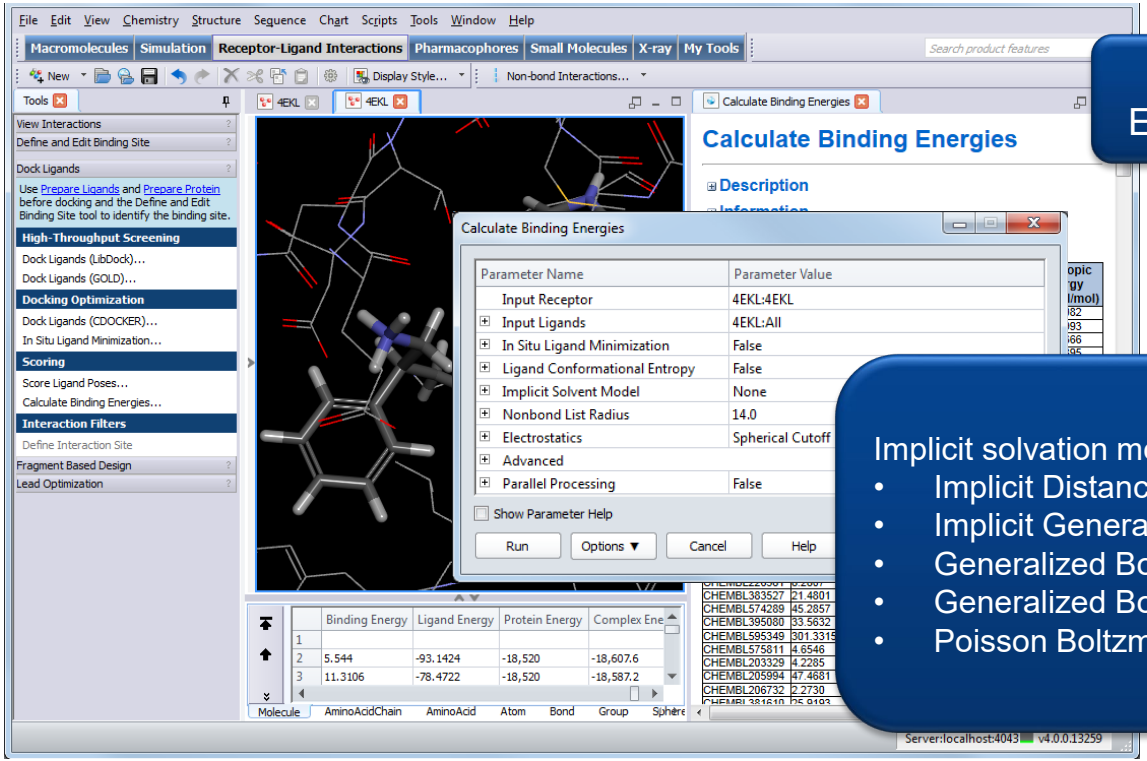
Search and filter for the hits

- Compare with experimental results
  - Key residues
  - RMSD with the reference structure
  - Salt bridge
- Calculate binding energy
- View the non-bond interactions between ligand and receptor
- View the non-bond interactions between ligand and receptor (2D diagram)

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Binding Energies



Binding energy

$$E_{binding} = E_{complex} - (E_{ligand} + E_{receptor})$$

Implicit solvation model can be used:

- Implicit Distance-Dependent Dielectrics
- Implicit Generalized Born
- Generalized Born with Molecular Volume(GBMV)
- Generalized Born with a Simple Switching(GBSW)
- Poisson Boltzmann with non-polar Surface Area (PBSA)

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Analysis - Analyze Ligand Poses

	Favorable Total	Unfavorable Total	Charge Total	Halogen Total	Hydrophobic Total	HydrogenBond Total	Other Total	Favorable A:LEU156	Favorable A:GLY157
1	19	0	4	0	9	5	1	0	0
2	17	0	2	0	3	10	2	0	0
3	16	0	0	0	10	5	1	0	0
4	15								
5	20								
6	14								
7	18								
8	12								
9	13								

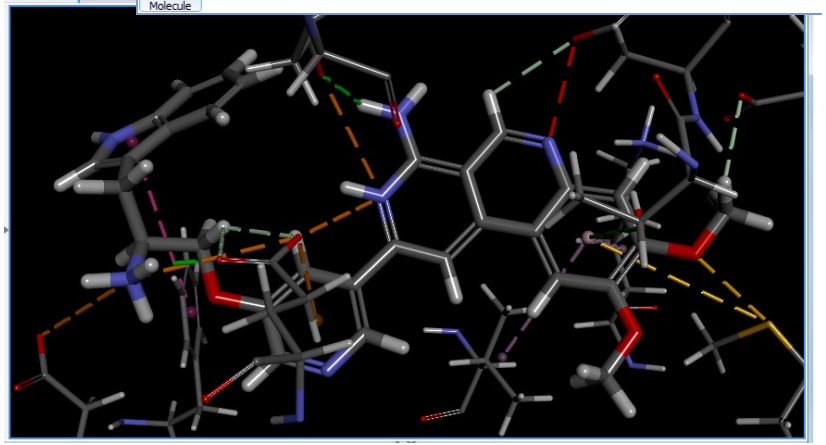
Molecule

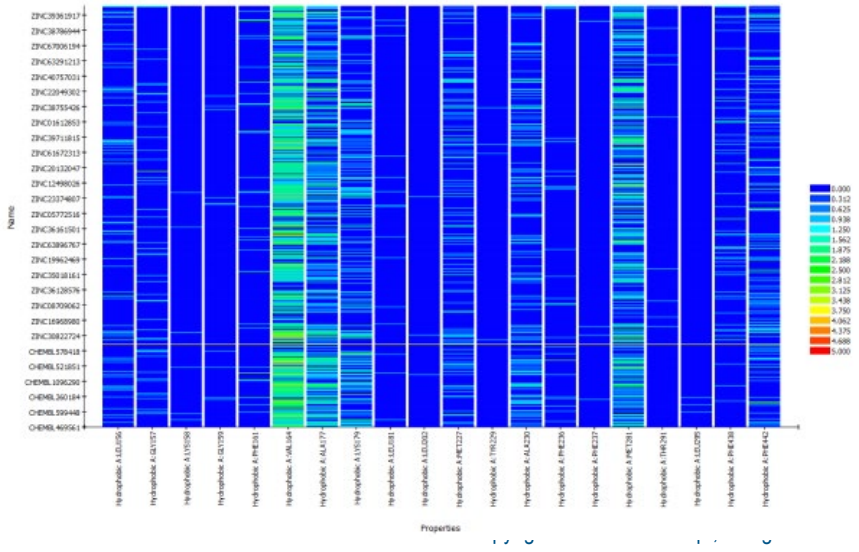
ptor-Ligand Interaction

4EKL

	Hydrophobic A:PHE438	Hydrophobic A:PHE442	HydrogenBond A:LEU156	HydrogenBond A:GLY157	HydrogenBond A:LYS158	HydrogenBond A:GLY159	HydrogenBond
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	1	0	0	0	0	0
4	0	0	0	0	1	0	0
5	0	0	0	0	2	0	0
6	0	1	0	0			
7	0	1	0	0			
8	0	0	0	1			
9	0	0	1	0			

Molecule

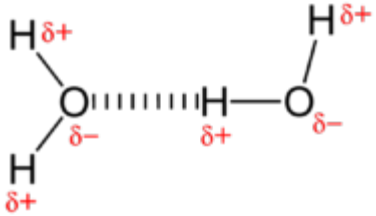


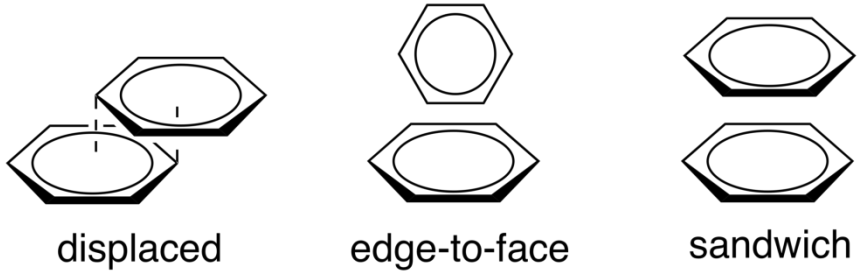


3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Non-covalent (Non-bond) interactions

- Hydrogen bond interactions
- Electrostatic interactions
- π-effects
- Van der waals forces
- Hydrophobic effects





3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

bonds strenghts

Covalent bond

Type	Energy
C-O bond	81 kcal/mol
C-C bond	86 kcal/mol
C-H bond	103 kcal/mol
C=C bond	143 kcal/mol
C=O bond	165 kcal/mol

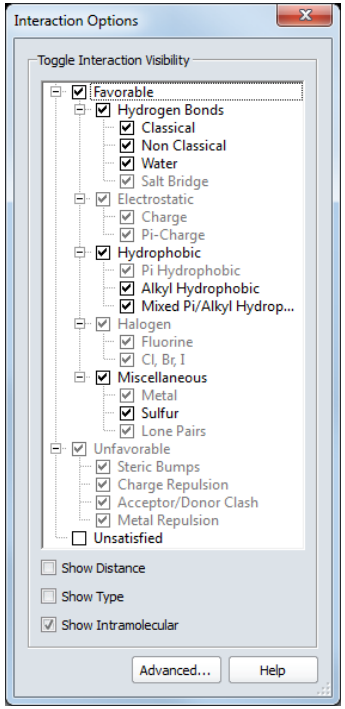
Non-Covalent bond

Type	Energy
Hydrophobic	<10 kcal/mol
Hydrogen bond	2-30 kcal/mol
Electrostatic	1-20 kcal/mol
π-π aromatic stacking	0-10 kcal/mol
Van der Waals	0.1-1 kcal/mol

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Analysis - Non-Bond Interactions



Favorable (*See below*)

Unfavorable
- Steric Bumps
- Charge Repulsion
- Acceptor-Acceptor clashes
- Donor-Donor clashes

Unsatisfied
- Hydrogen bond donor
- Hydrogen bond acceptor
- Charged atoms

- Charge
  - Attractive Charges
  - Salt Bridge
  - Pi-Cation
  - Pi-Anion
- Halogen
  - Halogen (Fluorine)
  - Halogen (Cl, Br, I)

- Hydrophobic
  - Pi-Pi Stacked
  - Pi-Pi T-Shaped
  - Amide-Pi Stacked
  - Alkyl
  - Pi-Sigma
  - Pi-Alkyl
- Other
  - Metal-Acceptor
  - Pi-Sulfur
  - Sulfur-X
  - Pi-Lone Pair

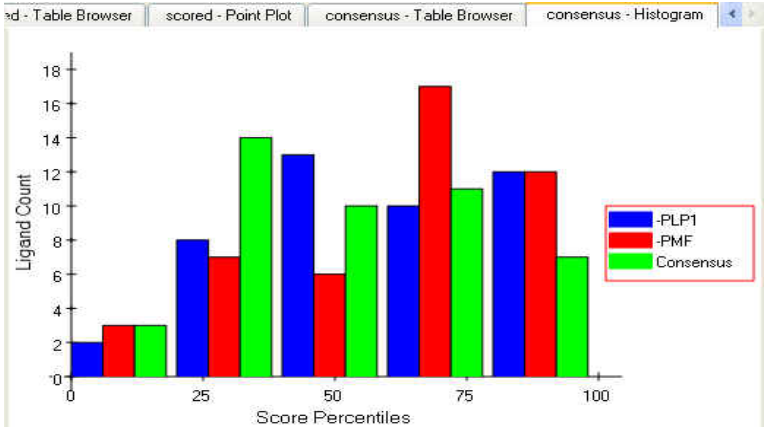
- Hydrogen Bond
  - Conventional Hydrogen Bond
  - Carbon Hydrogen Bond
  - Pi Donor Hydrogen Bond
  - Water Mediated Hydrogen Bond
  - Water Hydrogen Bond
  - Salt Bridge

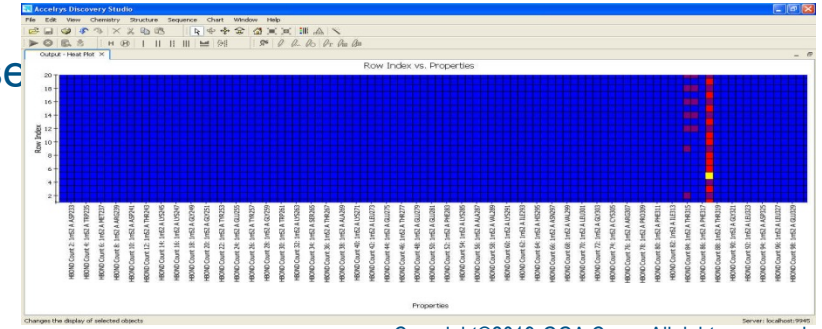
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

rights reserved.

Pose Analysis

- Plot Charts of Scoring
  - Simple Line Plot
  - Histogram
  - Hit Rate Plot
- Run Analyze Ligand Poses protocol
  - Hbonds
  - Contacts
  - Heat Maps
- Optional energy minimization of each pose
  - Ligand Minimization Protocol
- Fits saved to a molecule table
  - Can be exported to an SD file



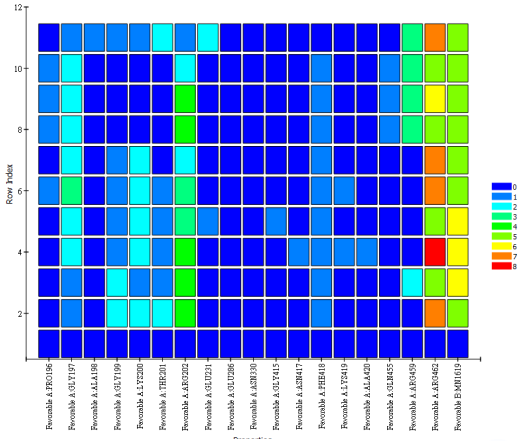


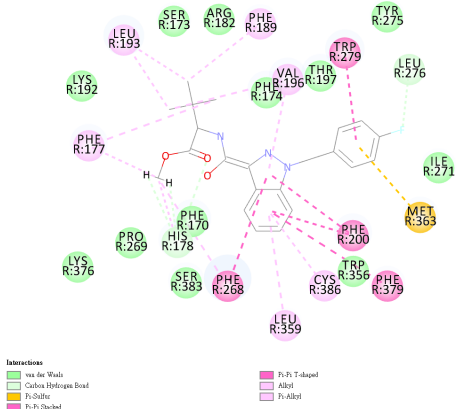
3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

Hands on

- Analyze FUB poses by Analyze Ligand Pose and calculate binding energies protocol
- Draw the heat map by Chart tools
- Find the poses which interact with
  - M363, W279
  - Generate 2D diagram and 3D interaction scheme

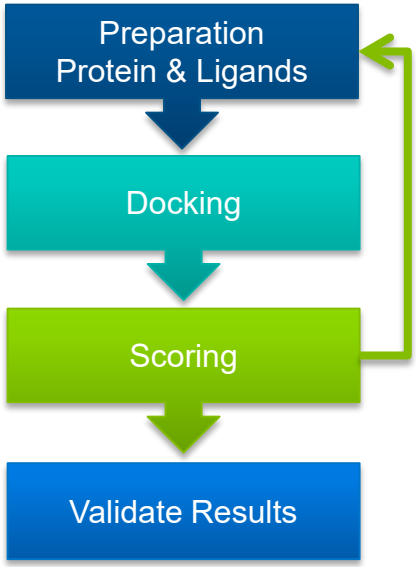




Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Docking Workflow



Repeat Cycle = Virtual High Throughput Screening (vHTS)

- In order to maximize the rate of success, the evaluation phase in each steps is very important.
- Requirements
  - Known active compounds and decoy
  - Bound molecules
  - Optimize the parameters of docking technique
  - Prioritized the molecules in library before proceeding the docking and scoring steps

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Q&A

Copyright©2019 GGA Corp., All rights reserved.

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |



創源生技
GGA

For more information please contact…



創源生技
GGA

台北市114內湖區新湖一路36巷28號

中華民國 台灣

Ph: (02) 2795 1777 x 3014 Fax: (02) 2793 8009

msc-support@gga.asia

www.gga.asia

3DS.COM/BIOVIA © Dassault Systèmes | Confidential Information | 3/16/2019 |

Copyright©2019 GGA Corp., All rights reserved.