Prediction of framework-guest systems using molecular docking

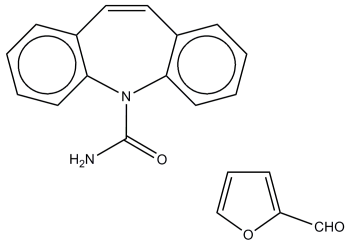
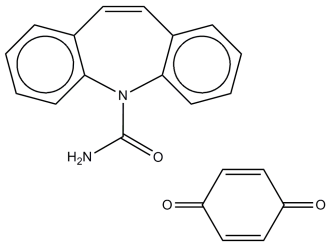
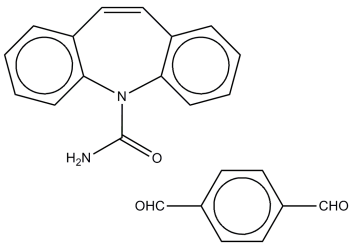
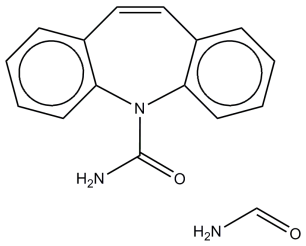
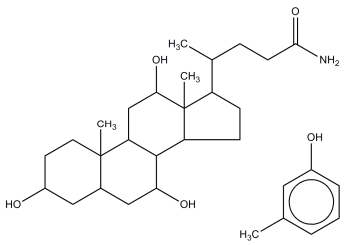
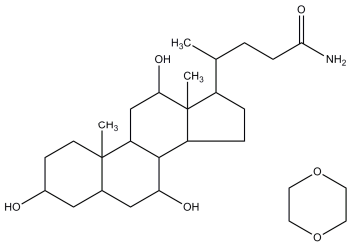
Oliver Korb and Peter A. Wood*

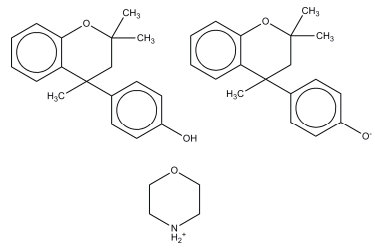
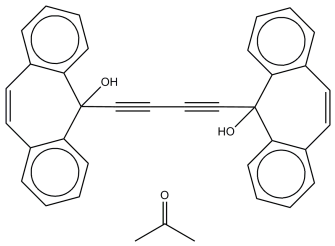
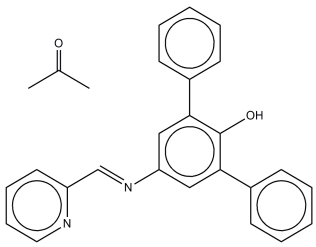
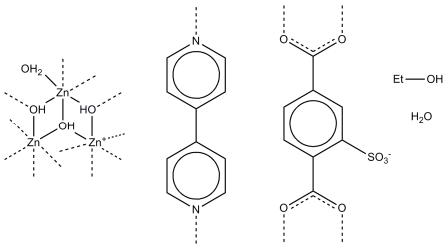
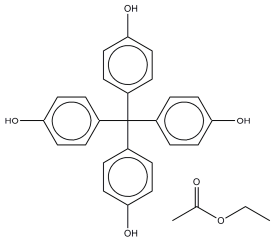
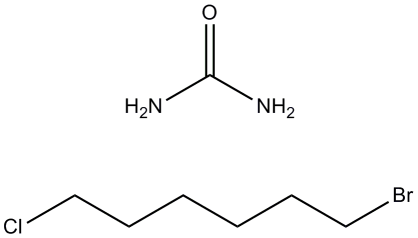
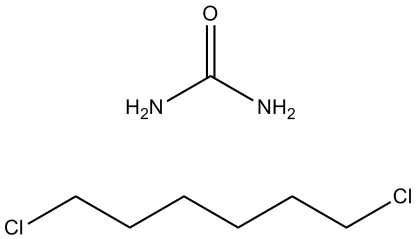
Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK.

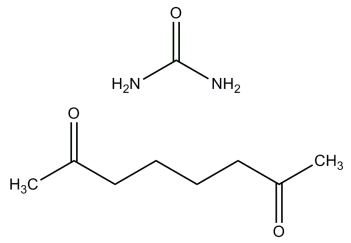
Supporting Information

Table of Contents

1. Host-Guest Framework Details

System Label	CSD Refcode*	Chemical name	Chemical Diagram
cbz1	FOMXAH[1]	carbamazepine:furfural co-crystal	
cbz2	UNEYOB[2]	carbamazepine:benzoquinone co-crystal	
cbz3	UNEYUH[2]	carbamazepine:terephthalaldehyde co-crystal	
cbz4	UNIBOI[2]	carbamazepine:formamidine co-crystal	
chol1	GOVQAJ[3]	cholamide *m*-cresol clathrate	
chol2	LAHGAC01[3]	cholamide 1,4-dioxane	

dian1	GASMOD[4]	Dianin's compound morpholine clathrate	
diol1	GIVCIY[5]	1,4-di-(5H-dibenzo[*a*,*d*]cyclohepten-5-ol)-buta-1,3-diyne acetone clathrate	
imin1	727852[6]	2-pyridyl 3,5-diphenyl-4-iminophenol acetone clathrate	
mof1	JITPOS[7]	catena-(bis(μ_4-2-sulfonylterephthalato)-hexakis(μ_3-hydroxo)-tris(μ_2-4,4'-bipyridine)-diaqua-hexa-zinc(II) diethanol tetrahydrate clathrate	
phen1	UJOFEE[8]	tetrakis(4-hydroxyphenyl)methane ethyl acetate clathrate	
urea1	WOMGOU[9]	urea 1-bromo-6-chlorohexane clathrate	
urea2	WOMGIO[9]	urea 1,6-dichlorohexane clathrate	

urea3	TOZHOF[10]	urea 2,7-octanedione clathrate	

*If CSD refcode is unavailable at time of publication, the deposition number is shown instead.

2. GOLD Configuration Files

- *gold.params* - GOLD parameter file
- *gold.conf* - GOLD configuration file (host-guest system and binding site definition)
- *chemplp_nocho.params* - CHEMPLP scoring function parameters (C-H...O interactions disabled)

N.B. The specific option used to fix the host framework system is not available in the current version of GOLD, but will be made available in the next release of the software.

(i) Changes to GOLD parameter file "gold.params"

```
RELAX_LIGAND = 1
MAX_RELAX_DIST = 2.0
MAX_RELAX_ANGLE = 60.0
```

(ii) GOLD configuration file "gold.conf"

```
  GOLD CONFIGURATION FILE

autoscale = 1.0

  POPULATION
popsiz = auto
select_pressure = auto
n_islands = auto
maxops = auto
niche_siz = auto

  GENETIC OPERATORS
pt_crosswt = auto
allele_mutatewt = auto
migratewt = auto

  FLOOD FILL
radius = 12
origin = 0 0 0
do_cavity = 0
floodfill_atom_no = 0

  DATA FILES
param_file = gold.params
set_ligand_atom_types = 1
set_protein_atom_types = 1
directory = .
tordist_file = DEFAULT
make_subdirs = 0
save_lone_pairs = 0
```

```
fit_points_file = fit_pts.mol2
read_fitpts = 0
output_file_format = MOL2
  FLAGS
internal_ligand_h_bonds = 0
n_ligand_bumps = 0
flip_free_corners = 0
flip_amide_bonds = 0
flip_planar_n = 1 flip_ring_NRR flip_ring_NHR
flip_pyramidal_n = 0
rotate_carboxylic_oh = flip
use_tordist = 1
postprocess_bonds = 1
rotatable_bond_override_file = DEFAULT
  TERMINATION
early_termination = 0
n_top_solutions = 3
rms_tolerance = 1.5

  CONSTRAINTS
force_constraints = 0

  COVALENT BONDING
covalent = 0

  SAVE OPTIONS
save_score_in_file = 1
save_protein_torsions = 1
clean_up_option save_top_n_solutions 1
clean_up_option delete_redundant_log_files
clean_up_option delete_empty_directories

  FITNESS FUNCTION SETTINGS
initial_virtual_pt_match_max = 6
relative_ligand_energy = 0

solvate_all = 1

# use scoring function CHEMPLP (CH..O interactions disabled)
gold_fitfunc_path = plp
score_param_file = chemplp_nocho.params

# fix host system (no rotatable donor groups)
fix_all_protein_rotatable_bonds = 1

# host system file
protein_datafile = host.mol2

# guest conformation used for docking (CORINA-generated structure)
ligand_data_file guest_corina.mol2 10

# guest conformation used for rmsd calculations (crystal structure)
ligand_reference_file = guest_crystal.mol2.

# include all host system atoms within 5 A from any guest heavy atom (crystal structure conformation)
cavity_file = guest_crystal.mol2
floodfill_center = cavity_from_ligand 5.0 atom
```

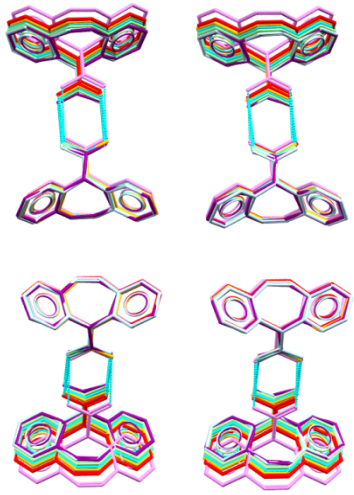
(iii) Parameter file "chemplp_nocho.params" neglecting CHO interactions for CHEMPLP

```
SCORING_FUNCTION PLP
# PLP coefficients
PLP_COEFFICIENT   1.0
PLP_LIGAND_CLASH_COEFFICIENT 1.0
PLP_LIGAND_TORSION_COEFFICIENT 2.0
PLP_PROTEIN_ENERGY_COEFFICIENT 1.0
PLP_CONSTRAINT_COEFFICIENT 1.0
PLP_WATER_BARRIER 3.0
PLP_GRID_SPACING 0.3
PLP_HBOND_METAL_FUNCTION CHEMSCORE
# ChemScore coefficients:
HBOND_COEFFICIENT -3.00
METAL_COEFFICIENT -6.00
CHARGED_HBOND_FACTOR 2.0
CHARGED_METAL_FACTOR 2.0
DELTA_BETA_IDEAL   80.0
DELTA_BETA_MAX     100.0
CHO_COEFFICIENT      -3.00
CHO_TYPE              OFF
CHO_R_IDEAL           2.35
CHO_DELTA_R_IDEAL       0.25
CHO_DELTA_R_MAX        0.65
CHO_ALPHA_IDEAL       180.0
CHO_DELTA_ALPHA_IDEAL  50.0
CHO_DELTA_ALPHA_MAX   100.0
CHO_BETA_IDEAL        180.0
CHO_DELTA_BETA_IDEAL   80.0
CHO_DELTA_BETA_MAX    100.0
HBOND_SCALING 0
# potential parameters:
HBOND_A 2.3
HBOND_B 2.6
HBOND_C 3.1
HBOND_D 3.4
HBOND_E -1.0
HBOND_F 20.0
BURIED_A 3.4
BURIED_B 3.6
BURIED_C 4.5
BURIED_D 5.5
BURIED_E -0.1
BURIED_F 20.0
NONPOLAR_A 3.4
NONPOLAR_B 3.6
NONPOLAR_C 4.5
NONPOLAR_D 5.5
NONPOLAR_E -0.4
NONPOLAR_F 20.0
METAL_A 1.4
METAL_B 2.2
METAL_C 2.6
METAL_D 2.8
METAL_E -1.0
METAL_F 20.0
REPULSIVE_A 3.2
REPULSIVE_B 5.0
```

REPULSIVE_C 0.1
REPULSIVE_D 20.0

3. Carbamazepine Framework Rigidity

Overlay diagram of the seven isostructural channel-based carbamazepine co-crystal structures showing the relative rigidity of the framework. Diagram generated using Mercury.[11] Identification of isostructural systems taken from recent paper by Childs and co-workers.[12]



4. Supporting Information References

[1] A. Johnston, A. J. Florence, A. R. Kennedy, *Acta Crystallographica, Section E* **2005**, *61*, o1777-o1779.

[2] S. G. Fleischman, S. S. Kuduva, J. A. McMahon, B. Moulton, R. D. B. Walsh, N. Rodriguez-Hornedo, M. J. Zaworotko, *Crystal Growth & Design* **2003**, *3*, 909-919.

[3] K. Sada, T. Kondo, M. Ushioda, Y. Matsuura, K. Nakano, M. Miyata, K. Miki, *Bulletin of the Chemical Society of Japan* **1998**, *71*, 1931-1937.

[4] G. O. Lloyd, M. W. Bredenkamp, L. J. Barbour, *Chemical Communications* **2005**, 4053-4055.

[5] A. Jacobs, N. L. Z. Masuku, L. R. Nassimbeni, J. H. Taljaard, *CrystEngComm* **2008**, *10*, 322-326.

[6] A. Bacchi, M. Carchelli, T. Chiodo, F. Mezzadri, F. Nestola, A. Rossi, *Crystal Growth & Design* **2009**, *9*, 3749-3758.

[7] S. Horike, S. Bureekaew, S. Kitagawa, *Chemical Communications* **2008**, 471-473.

[8] J.-H. Fournier, T. Maris, M. Simard, J. D. Wuest, *Crystal Growth & Design* **2003**, *3*, 535-540.

[9] M. D. Hollingsworth, U. Werner-Zwanziger, M. E. Brown, J. D. Chaney, J. C. Huffman, K. D. M. Harris, S. P. Smart, *Journal of the American Chemical Society* **1999**, *121*, 9732-9733.

[10] M. E. Brown, J. D. Chaney, B. D. Santarsiero, M. D. Hollingsworth, *Chemistry of Materials* **1996**, *8*, 1588-1591.

[11] C. F. Macrae, et al., *Journal of Applied Crystallography* **2008**, *41*, 466-470.

[12] S. L. Childs, P. A. Wood, N. Rodriguez-Hornedo, L. S. Reddy, K. I. Hardcastle, *Crystal Growth & Design* **2009**, *9*, 1869–1888.