Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2014

A catalytic role for methionine revealed by a combination of computation and experiments on phosphite dehydrogenase

ELECTRONIC SUPPLEMENTARY INFORMATION

Kara E. Ranaghan[a,b], John E. Hung[c], Gail J. Bartlett[b], Tiddo J. Mooibroek[b], Jerremy N. Harvey[a,b], Derek N. Woolfson[b,d], Wilfred A. van der Donk[c]*, and Adrian J. Mulholland[a,b]*

AUTHOR ADDRESS

[a]Centre for Computational Chemistry,

[b]School of Chemistry,

University of Bristol,

Cantock's Close,

Clifton,

Bristol BS8 1TS.

[c]Department of Chemistry and the Howard Hughes Medical Institute,

University of Illinois at Urbana-Champaign,

600 South Mathews Avenue,

Urbana, Illinois 61801.

[d]School of Biochemistry,

Medical Sciences,

University of Bristol,

Clifton,

Bristol BS8 1TD.

Contents

1. Materials

All mutants were generated in the 17X-PTDH background[1]. The 16 total mutations in TS-PTDH compared to wt-PTDH from *P. stutzeri* are D13E, M26I, V71I, E130K, Q132R, Q137R, I150F, Q215L, R275Q, L276Q, I313L, V315A, A319E, A325V, E332N, and C336D. The 17X mutant has one additional mutation (E175A) that allows the protein to use both NAD+ and NADP; it is this protein that is most versatile for cofactor regeneration [note that this enzyme is called 12X in references [1] and [2] but actually contains 17 mutations].

2. Preparation of PTDH Mutant Constructs.

The mutants were prepared by the polymerase chain reaction (PCR) using QuikChange mutagenesis (Agilent). The primers used to incorporate the mutations are provide below, along with their reverse complement strands, which are not shown: M53A: 5'-CG ATG ATG GCG TTC GCG CCC GAT CGG GTC G-3'; M53L: 5'-CG ATG ATG GCG TTC TTG CCC GAT CGG GTC– 3'; M53N: 5'-CAG GCG ATG ATG GCG TTC AAT CCC GAT CGG G-3'. The mutated codon is underlined in each case. Once the mutant genes were constructed, they were sequenced in their entirety to ensure that the desired mutation was incorporated and no other mutations were generated.

3. Overexpression and Purification of PTDH mutants.

Chemically competent *E. coli* BL21 (DE3) cells were transformed with plasmids containing the desired gene in pET15b and plated on a LB agar plate containing 100 µg/mL ampicillin. A single colony was picked and used to inoculate an overnight culture (50 mL). The cells were harvested by centrifugation, resuspended in 1 L of LB/ampicillin and grown at 37 °C to $OD_{600\ nm}$ = 0.6. IPTG was added to the culture (0.3 mM final concentration), which was incubated at 18 °C for 16-20 h. Cells were then pelleted by centrifugation (8,000xg, 30 min), resuspended in 20 mM Tris pH 7.6, 0.5 M NaCl, 10% glycerol (v/v), and stored at −80 °C. The frozen pellet was then thawed on ice and lysed using an Avestin Emulsiflex-C3 Homogenizer (7500 Psi,6 passes). The insoluble lysate was removed by centrifugation (15,000xg, 30 min), and the soluble lysate was passed through a 0.44 µm filter. The enzyme was purified via FPLC (Äkta) by immobilized metal affinity chromatography (IMAC) with a HisTrap HP Ni^{2+}-affinity column (GE Life Sciences). The column was washed with 5 column volumes (CV) of buffer A, followed by a gradient to 40% buffer B over 5 CV, and another gradient of 2 CV to 100% buffer B. (Buffer A: 20 mM Tris pH 7.6, 100 mM NaCl, 10 mM imidazole, 10% glycerol (v/v); buffer B: 20 mM Tris pH 7.6, 100 mM NaCl, 500 mM imidazole, 10% glycerol (v/v)). Fractions containing 17X-PTDH were identified by activity assay and analyzed by SDS-PAGE. The highly pure fractions were pooled and concentrated using centrifugal filtration (Amicon membrane, 30 kDa molecular weight cutoff; Millipore), followed by buffer exchange or dialysis to 20 mM MOPS pH 7.6, 100 mM KCl, 10% glycerol. The protein solution was then flash frozen and stored at −80 °C. Protein concentration was determined using the calculated extinction coefficient for 17X-PTDH, 28,000 $M^{-1}\,cm^{-1}$ [3]. All

buffers used for SeMet PTDH preparation were supplemented with 5 mM dithiothreitol (DTT) to prevent selenomethionine oxidation.

4. Preparation of QM/MM model

The model of PTDH used in QM/MM simulations was prepared based on the crystal structure 4E5K[4]. Missing residues, sequence mismatches and sidechain flips identified by WHATIF are given in Table S6. Manual inspection of the crystal structure and predictions made by the PROPKA[5] program were used to assign the protonation states of titratable residues in the enzyme. The results are given below. Special attention was paid to the protonation of histidine residues as they can be singly protonated on either ND1 or NE2, or doubly protonated. The CHARMM residue name HSE indicates that the residue is protonated at NE2, HSD indicates that the protonation is at HSD, and HSP indicated a doubly protonated residue.

Protonation of Histidines

RESIDUE	pKa	pKmodel	CHARMM RES	H-Bonds
HIS 9A	1.40	6.50	HSE	NE2 → THR 33 OG1
HIS 12A	8.36	6.50	HSE	
HIS 22A	7.01	6.50	HSD	ND1 → LEU 18/19 O
HIS 118A	4.62	6.50	HSE	NE2 → TYR 139 O / PHE 138 O
HIS 174A	9.09	6.50	HSD	ND1 → VAL 192 O
HIS 216A	4.05	6.50	HSE	NE2 → GLU 264 OE1
HIS 284A	6.29	6.50	HSD	
HIS 292A	4.55	6.50	HSP	ND1 → HIS 292 O
HIS 9B	2.28	6.50	HSD	ND1 → THR 8 OG1
HIS 12B	5.14	6.50	HSE	NE2 → ARG 298 O
HIS 22B	7.18	6.50	HSD	ND1 → LEU 18 O
HIS 118B	3.84	6.50	HSE	NE2 → TYR 139 O
HIS 174B	3.91	6.50	HSD	ND1 → VAL 192 O
HIS 216B	9.22	6.50	HSD	ND1 → SER 239 O
HIS 284B	6.36	6.50	HSE	
HIS 292B	4.57	6.50	HSP	ND1 → HIS 292 O
HIS 9C	1.95	6.50	HSD	ND1 → THR 8 OG1
HIS 12C	3.11	6.50	HSD	ND1 → GLU 305 OE1
HIS 22C	8.72	6.50	HSE	NE2 → GLU 67A OE2
HIS 118C	3.69	6.50	HSE	NE2 → TYR 139 O
HIS 174C	7.01	6.50	HSD	
HIS 216C	7.21	6.50	HSE	NE2 → GLU 264 OE2
HIS 284C	6.37	6.50	HSD	
HIS 292C	7.26	6.50	HSD	ND1 → GLU 266 OE1
HIS 9D	1.30	6.50	HSE	NE2 → THR 33 OG1
HIS 12D	7.34	6.50	HSE	*ND1 → GLU 305 OE1 in opt h-b
HIS 22D	6.98	6.50	HSD	ND1 → LEU 18/19 O
HIS 118D	4.16	6.50	HSE	NE2 → TYR 139 O
HIS 174D	9.23	6.50	HSE	*ND1 → VAL 192 O in opt h-b
HIS 216D	9.99	6.50	HSE	NE2 → GLU 264 OE1
HIS 284D	6.50	6.50	HSD	
HIS 292D	4.01	6.50	HSP	ND1 → GLU 266 OE2
HIS 9E	2.47	6.50	HSE	
HIS 12E	6.94	6.50	HSD	ND1 → GLU 305 OE1
HIS 22E	8.04	6.50	HSE	NE2 → GLU 67D OE2
HIS 118E	4.58	6.50	HSD	ND1 → PHE 138 O / TYR 139 O

```
HIS 174E   6.36      6.50      HSD ND1 → VAL 192 O
HIS 216E   7.29      6.50      HSE NE2 → GLU 264 OE1
HIS 284E   6.36      6.50      HSE
HIS 292E   4.90      6.50      HSD ND1 → GLU 266 OE2
HIS   9F   1.68      6.50      HSE
HIS  12F   1.19      6.50      HSD ND1 → GLU 305 OE1
HIS  22F   5.30      6.50      HSD *NE2 → GLU 67B OE1 in opt h-b
HIS 118F   4.61      6.50      HSE NE2 → TYR 139 O
HIS 174F   6.87      6.50      HSD ND1 → VAL 192 O
HIS 216F   7.23      6.50      HSE NE2 → GLU 264 OE1
HIS 284F   6.36      6.50      HSD
HIS 292F   7.26      6.50      HSD ND1 → GLU 266 OE1
```

pKa's of active site residues

```
RESIDUE    pKa   pKmodel      H-Bonds

ASP  79A   2.67      3.80      N → ASP 79 OD1
ASP  79B   2.50      3.80      N → ASP 79 OD1

ASP  79C   3.44      3.80      *not in optimal h-bond
ASP  79D   2.25      3.80      N → ASP 79 OD1

ASP  79E   4.92      3.80      *not in optimal h-bond
ASP  79F   1.26      3.80      N → ASP 79 OD1

GLU 266A   2.87      4.50      N → PHE 263 O
GLU 266B   3.26      4.50      N → PHE 263 O
GLU 266C   2.93      4.50      N → PHE 263 O
GLU 266D   3.72      4.50      N → GLU 243 OE2, N → PHE 263 O
GLU 266E   4.19      4.50      N → GLU 243 OE2, N → PHE 263 O
GLU 266F   2.93      4.50      N → GLU 243 OE2, N → PHE 263 O
HIS 292A   5.41      6.50      N → THR 290 O/ ASP 261 O ND1 → GLU266 OE2
HIS 292B   5.46      6.50      N → THR 290 O/ ASP 261 O ND1 → GLU266 OE2
HIS 292C   7.17      6.50      N → THR 290 O/ ASP 261 O ND1 → GLU266 OE1
HIS 292D   4.93      6.50      N → THR 290 O/ ASP 261 O ND1 → GLU266 OE2
HIS 292E   0.00      6.50      N → THR 290 O/ ASP 261 O ND1 → HIS292 O
HIS 292F   7.18      6.50      N → THR 290 O/ ASP 261 O ND1 → GLU266 OE1
ARG 237A  11.73     12.50      NE → ASP 79 OD1, NH1 → CYS 236 O, NH2 → GLY
77 O, NH2 → ASP 79 OD1
ARG 237B  11.93     12.50      N → CYS 236 SG, NE → ASP 79 OD1, NH1 → CYS
236 O, NH2 → MET  53 SD, NH2 → GLY 77

                               O, NH2 → ASP 79 OD1

ARG 237C  13.79     12.50      NH2 → ASP 267 OD1
ARG 237D  12.34     12.50      N→ ASN 234 OD1, N → CYS 236 SG, NE → ASP 79
OD1, NH1 → CYS 236 O, NH2 → ASP 79 OD1
ARG 237E  13.25     12.50      N → CYS 236 SG, NH1 → GLU 266 OE1
ARG 237F  15.09     12.50      N→ ASN 234 OD1, N → CYS 236 SG, NH1 → ASP
267 OD1, NH2 → ASP 79 OD1
ARG 301A  11.43     12.50      N → VAL 297 O, NH1 → ALA 74 O, NH1 SER 295 OG
ARG 301B  11.02     12.50      N → VAL 297 O, NH1 → ALA 74 O, NH1 SER 295 OG
ARG 301C  10.99     12.50      N → VAL 297 O, NH1 → ALA 74 O, NH1 SER 295 OG
ARG 301D  11.25     12.50      N → VAL 297 O, NH1 → ALA 74 O, NH1 SER 295
OG, NH2 → ALA 74 O
```

```
ARG 301E   10.64      12.50       N → VAL 297 O, NH1 → ALA 74 O, NH1 SER 295 OG
ARG 301F   10.79      12.50       N → VAL 297 O, NH1 → ALA 74 O, NH1 SER 295 OG
```

5. Analysis of structures from QM/MM simulations.

For the purposes of analyzing interactions in the active site of the enzyme, a hydrogen bond was defined as a hydrogen to acceptor distance of £≤ 2.6 Å (2.8 Å if the acceptor is sulfur) and a donor to acceptor distance of £≤ 3.5 Å (4.0 Å if the acceptor is sulfur), the angle between atoms forming the hydrogen bond must also be greater than 90°[6]. The contributions of different residues (e.g. to QM/MM electrostatic interactions) were calculated by individually deleting the residues (identified as important in the hydrogen bond analysis) from 110 snapshots in the dynamics trajectory in the reactant (R), transition state 1 (TS1), intermediate (Int), transition state 2 (TS2) and product complexes and recalculating the energy. The average electrostatic interaction energy is then the average of the difference between the original QM/MM electrostatic energy and the energy with the residue deleted. It is important to note that the interaction energies calculated in this manner correspond to complete deletion of a residue and also do not include the effects of dielectric shielding. As a result, the energies are not directly comparable to experimental results (e.g. $\Delta\Delta G^{\ddagger}$ for mutated enzymes); however, this type of decomposition analysis is useful for identifying important interactions in enzyme-catalyzed reactions, e.g. identifying residues that have the largest effects on reaction barriers and reaction energies[7-10].

6. Analysis of the interaction between Met53 and the NAD+ co-factor

Optimized model complexes representing the interaction of methionine with the nicotinamide ring of the NAD^+ cofactor did not show an interaction through the methyl group of methionine. Instead, a face-on complex between the sulfur of methylthioethane and the aromatic ring of the nicotinamide fragment was observed, with a similar interaction energy to that found for the face-on complex between the 4-methylimidazolium and methylthioethane [−11.9 kcal/mol nicotinamide / methylthioethane complex; −10.4 kcal/mol 4-methylimidazolium / methylthioethane complex (MP2/aug-ccVTZ level with counterpoise correction for BSSE)]. Although a C-H···O hydrogen bond between the methyl and amide of the methionine and nicotinamide fragments is not a minimum energy structure for the model complex, this interaction is likely to be favorable, though weak. Recent NMR experiments have shown similar coordination of the methyl group of S-adenosylmethionine through C−H−O hydrogen bonds to tyrosine and glycine residues in the active site of lysine methyltransferase SET7/9[11].

7. NBO analysis of the face-on complex

Natural bond order (NBO) analysis[12] (B3LYP/6-311+G(d,p)) of the model complex was carried out using Gaussian NBO version 3.1 with the RESONANCE keyword to allow for delocalized orbitals in aromatic systems. NBO analysis shows that the lone pairs of electrons on sulfur interact with several antibonding orbitals on the 4-methylimidazolium fragment, confirming that the interaction between the two fragments can in part be described as an n→π* interaction. However, second order

perturbation theory analysis of the Fock matrix in the NBO basis indicates that the energy benefit from the favorable overlap is small (< 1 kcal/mol). Below is the summary of the NBO analysis and the second order perturbation theory analysis of the Fock matrix in the NBO basis. Full output of the calculation is available on request.

```
Natural Bond Orbitals (Summary):

                                                                    Principal
Delocalizations
           NBO                          Occupancy    Energy
(geminal,vicinal,remote)

================================================================================
=============
 Molecular unit  1  (C3H8S)
   1. BD (   1) C   1 - H   2          1.98879    -0.59266
297(v),73(v)
   2. BD (   1) C   1 - H   3          1.98819    -0.59482  298(v)
   3. BD (   1) C   1 - C   4          1.99520    -0.73121  298(g)
   4. BD (   1) C   1 - H  25          1.97971    -0.58972  299(v)
   5. BD (   1) C   4 - H   5          1.98523    -0.61341
293(v),98(v)
   6. BD (   1) C   4 - H   6          1.97931    -0.61383
294(v),300(v),99(v),45(v)
   7. BD (   1) C   4 - S   7          1.98326    -0.70489
296(v),303(v),119(v),44(v)
   8. BD (   1) S   7 - C   8          1.99023    -0.71579
298(v),71(v)
   9. BD (   1) C   8 - H   9          1.99543    -0.62961
  10. BD (   1) C   8 - H  10          1.99479    -0.62898  98(v)
  11. BD (   1) C   8 - H  11          1.98960    -0.63098
299(v),99(v)
  27. CR (   1) C   1                  1.99935   -10.11756  72(v),61(v)
  28. CR (   1) C   4                  1.99925   -10.17657
299(g),47(v)
  29. CR (   1) S   7                  2.00000   -87.86205
  30. CR (   2) S   7                  1.99943    -8.88890  122(v)
  31. CR (   3) S   7                  1.99992    -5.98993
  32. CR (   4) S   7                  1.99994    -5.98909
  33. CR (   5) S   7                  1.99995    -5.98814
  34. CR (   1) C   8                  1.99928   -10.17388
300(g),137(v)
  41. LP (   1) S   7                  1.98315    -0.72764
295(v),302(v),297(v),313(r)
                                                   316(r),307(r)
  42. LP (   2) S   7                  1.93184    -0.34540
301(v),295(v),302(v),297(v)

296(r),120(v),72(v),313(r)

316(r),183(r),185(r),250(r)
                                                   307(r),231(r)
  44. RY*(   1) C   1                  0.00297     0.55131
```

45.	RY*(2)	C	1	0.00178	0.51274
46.	RY*(3)	C	1	0.00056	0.79517
47.	RY*(4)	C	1	0.00043	0.98614
48.	RY*(5)	C	1	0.00017	1.16345
49.	RY*(6)	C	1	0.00013	0.48239
50.	RY*(7)	C	1	0.00004	1.67914
51.	RY*(8)	C	1	0.00000	22.98286
52.	RY*(9)	C	1	0.00001	2.19276
53.	RY*(10)	C	1	0.00000	1.80234
54.	RY*(11)	C	1	0.00001	0.78622
55.	RY*(12)	C	1	0.00000	2.82349
56.	RY*(13)	C	1	0.00000	1.54133
57.	RY*(14)	C	1	0.00001	1.87919
58.	RY*(15)	C	1	0.00000	1.73889
59.	RY*(16)	C	1	0.00000	1.30312
60.	RY*(17)	C	1	0.00000	1.84624
61.	RY*(1)	H	2	0.00112	0.96327
62.	RY*(2)	H	2	0.00014	1.58586
63.	RY*(3)	H	2	0.00010	1.67872
64.	RY*(4)	H	2	0.00008	1.64703
65.	RY*(5)	H	2	0.00003	2.57523
66.	RY*(1)	H	3	0.00067	0.69384
67.	RY*(2)	H	3	0.00016	1.70847
68.	RY*(3)	H	3	0.00010	1.66516
69.	RY*(4)	H	3	0.00006	1.85744
70.	RY*(5)	H	3	0.00003	2.55810
71.	RY*(1)	C	4	0.00322	0.71612
72.	RY*(2)	C	4	0.00263	0.85977
73.	RY*(3)	C	4	0.00120	0.94211
74.	RY*(4)	C	4	0.00111	0.87615
75.	RY*(5)	C	4	0.00044	1.16503
76.	RY*(6)	C	4	0.00035	1.47466
77.	RY*(7)	C	4	0.00017	0.91917
78.	RY*(8)	C	4	0.00009	0.91671
79.	RY*(9)	C	4	0.00006	0.75358
80.	RY*(10)	C	4	0.00003	1.32470
81.	RY*(11)	C	4	0.00001	1.79649
82.	RY*(12)	C	4	0.00000	2.52747
83.	RY*(13)	C	4	0.00001	1.26846
84.	RY*(14)	C	4	0.00000	2.21083
85.	RY*(15)	C	4	0.00000	1.76782
86.	RY*(16)	C	4	0.00000	2.49357
87.	RY*(17)	C	4	0.00000	22.89356
88.	RY*(1)	H	5	0.00220	0.57405
89.	RY*(2)	H	5	0.00016	1.87916
90.	RY*(3)	H	5	0.00011	1.85226
91.	RY*(4)	H	5	0.00006	1.99837
92.	RY*(5)	H	5	0.00002	2.24586
93.	RY*(1)	H	6	0.00097	0.61115
94.	RY*(2)	H	6	0.00018	1.88901
95.	RY*(3)	H	6	0.00012	1.77947
96.	RY*(4)	H	6	0.00006	1.85529
97.	RY*(5)	H	6	0.00003	2.31712
98.	RY*(1)	S	7	0.00585	0.42462
99.	RY*(2)	S	7	0.00466	0.58200
100.	RY*(3)	S	7	0.00285	0.45861
101.	RY*(4)	S	7	0.00151	0.70816
102.	RY*(5)	S	7	0.00130	0.62785

103.	RY*(	6)	S	7				0.00067	0.45184
104.	RY*(	7)	S	7				0.00060	0.57697
105.	RY*(	8)	S	7				0.00042	0.31626
106.	RY*(	9)	S	7				0.00033	0.45914
107.	RY*(	10)	S	7				0.00025	0.69465
108.	RY*(	11)	S	7				0.00011	1.03103
109.	RY*(	12)	S	7				0.00008	1.69662
110.	RY*(	13)	S	7				0.00004	1.82742
111.	RY*(	14)	S	7				0.00002	1.51201
112.	RY*(	15)	S	7				0.00002	2.16911
113.	RY*(	16)	S	7				0.00002	0.91948
114.	RY*(	17)	S	7				0.00000	16.67964
115.	RY*(	18)	S	7				0.00000	182.20681
116.	RY*(	19)	S	7				0.00000	13.53121
117.	RY*(	20)	S	7				0.00000	16.81868
118.	RY*(	21)	S	7				0.00000	16.42442
119.	RY*(	1)	C	8				0.00313	0.71217
120.	RY*(	2)	C	8				0.00147	1.28494
121.	RY*(	3)	C	8				0.00059	0.80468
122.	RY*(	4)	C	8				0.00031	0.38237
123.	RY*(	5)	C	8				0.00019	0.96909
124.	RY*(	6)	C	8				0.00014	0.54064
125.	RY*(	7)	C	8				0.00006	0.53480
126.	RY*(	8)	C	8				0.00004	1.83467
127.	RY*(	9)	C	8				0.00002	1.06317
128.	RY*(	10)	C	8				0.00003	0.57203
129.	RY*(	11)	C	8				0.00000	2.71735
130.	RY*(	12)	C	8				0.00001	1.83203
131.	RY*(	13)	C	8				0.00000	2.38349
132.	RY*(	14)	C	8				0.00000	2.06106
133.	RY*(	15)	C	8				0.00000	1.49295
134.	RY*(	16)	C	8				0.00000	1.79545
135.	RY*(	17)	C	8				0.00000	23.14470
136.	RY*(	1)	H	9				0.00163	0.48419
137.	RY*(	2)	H	9				0.00016	2.09821
138.	RY*(	3)	H	9				0.00011	1.78232
139.	RY*(	4)	H	9				0.00006	1.81719
140.	RY*(	5)	H	9				0.00001	2.22660
141.	RY*(	1)	H	10				0.00200	0.52499
142.	RY*(	2)	H	10				0.00016	2.01058
143.	RY*(	3)	H	10				0.00008	1.71950
144.	RY*(	4)	H	10				0.00005	1.72645
145.	RY*(	5)	H	10				0.00003	2.37048
146.	RY*(	1)	H	11				0.00085	0.57255
147.	RY*(	2)	H	11				0.00014	1.97699
148.	RY*(	3)	H	11				0.00010	1.73690
149.	RY*(	4)	H	11				0.00006	1.57602
150.	RY*(	5)	H	11				0.00002	2.40629
288.	RY*(	1)	H	25				0.00076	0.73170
289.	RY*(	2)	H	25				0.00014	1.62885
290.	RY*(	3)	H	25				0.00008	1.66475
291.	RY*(	4)	H	25				0.00005	1.81247
292.	RY*(	5)	H	25				0.00003	2.59147
293.	BD*(	1)	C	1	-	H	2	0.00805	0.30677
294.	BD*(	1)	C	1	-	H	3	0.00832	0.30750
295.	BD*(	1)	C	1	-	C	4	0.01487	0.28928
296.	BD*(	1)	C	1	-	H	25	0.00823	0.30816
297.	BD*(	1)	C	4	-	H	5	0.02337	0.28689

```
 298. BD*(   1) C   4 - H   6          0.01481     0.28338
 299. BD*(   1) C   4 - S   7          0.02759     0.06479
 300. BD*(   1) S   7 - C   8          0.01170     0.06370
 301. BD*(   1) C   8 - H   9          0.01790     0.28633
 302. BD*(   1) C   8 - H  10          0.01645     0.28086
 303. BD*(   1) C   8 - H  11          0.00578     0.27589
       -------------------------------
              Total Lewis   41.78185  ( 99.4995%)
        Valence non-Lewis    0.15705  (  0.3740%)
        Rydberg non-Lewis    0.05313  (  0.1265%)
       -------------------------------
            Total unit  1   41.99203  (100.0000%)
           Charge unit  1    0.00797

 Molecular unit  2  (C4H7N2)
  12. BD (   1) H  12 - N  22          1.98820    -0.88205
311(v),312(v),227(v),266(v)
  13. BD (   1) C  13 - H  14          1.97449    -0.65431
313(v),312(v),212(v)
  14. BD (   1) C  13 - H  15          1.98007    -0.65251  310(v)
  15. BD (   1) C  13 - H  16          1.97128    -0.65172
313(v),312(v)
  16. BD (   1) C  13 - C  19          1.98448    -0.82811
312(g),311(v),266(v),188(v)
                                                   317(v),310(g)
  17. BD (   1) N  17 - H  18          1.98699    -0.87671
312(v),210(v),315(v),227(v)
                                                   310(g)
  18. BD (   1) N  17 - C  19          1.98386    -1.00414
318(v),314(v),311(g),230(v)

315(v),317(v),308(g),309(g)
                                                   312(g),306(v)
  19. BD (   1) N  17 - C  20          1.98704    -1.04149
308(v),304(v),210(v),310(g)

211(v),212(v),309(g),314(g)
  20. BD (   1) C  19 - C  23          1.97599    -0.90135
309(v),304(v),308(g),318(g)

156(v),249(v),252(v),317(g)
                                                   188(v)
  21. BD (   2) C  19 - C  23          1.79546    -0.44839
316(v),307(v),305(v),189(v)

250(v),157(v),299(r)
  22. BD (   1) C  20 - H  21          1.98603    -0.76735
310(v),317(v),249(v),188(v)
  23. BD (   1) C  20 - N  22          1.98830    -1.04624
309(v),318(v),266(v),317(g)

268(v),304(g),190(v),314(g)
  24. BD (   2) C  20 - N  22          1.92150    -0.54988
313(v),267(v)
  25. BD (   1) N  22 - C  23          1.98481    -1.00222
308(v),314(v),315(g),230(v)

311(v),310(v),312(g),304(g)
```

```
                                                                 212(v),211(v)
  26. BD (   1) C  23 - H  24          1.98226    -0.72675
315(v),312(g),310(v),210(v)
                                                                 249(v)
  35. CR (   1) C  13                  1.99927   -10.18459
210(v),312(v),183(v),310(v)
                                                                 178(v)
  36. CR (   1) N  17                  1.99928   -14.39074
211(v),212(v),230(v),205(v)
  37. CR (   1) C  19                  1.99918   -10.26707
158(v),312(g),311(v),318(v)

266(v),268(v),190(v)
  38. CR (   1) C  20                  1.99940   -10.32189
310(v),317(v),252(v),244(v)
  39. CR (   1) N  22                  1.99936   -14.39794
268(v),230(v),151(v)
  40. CR (   1) C  23                  1.99907   -10.25366
210(v),308(v),312(g),315(v)
                                                                 283(v),211(v)
  43. LP (   1) N  17                  1.51525    -0.45781
316(v),313(v),229(v),211(v)

206(v),214(v),189(g),212(v)
                                                                 300(r)
 151. RY*(   1) H  12                  0.00046     2.31216
 152. RY*(   2) H  12                  0.00042     1.33117
 153. RY*(   3) H  12                  0.00028     0.80233
 154. RY*(   4) H  12                  0.00015     1.73248
 155. RY*(   5) H  12                  0.00004     2.19442
 156. RY*(   1) C  13                  0.00234     0.51360
 157. RY*(   2) C  13                  0.00112     0.58045
 158. RY*(   3) C  13                  0.00056     0.93767
 159. RY*(   4) C  13                  0.00017     0.26529
 160. RY*(   5) C  13                  0.00016     0.51004
 161. RY*(   6) C  13                  0.00012     2.31238
 162. RY*(   7) C  13                  0.00005     0.52787
 163. RY*(   8) C  13                  0.00002     1.04275
 164. RY*(   9) C  13                  0.00003     1.14603
 165. RY*(  10) C  13                  0.00001     1.08369
 166. RY*(  11) C  13                  0.00000    22.84187
 167. RY*(  12) C  13                  0.00000     2.79118
 168. RY*(  13) C  13                  0.00000     2.96436
 169. RY*(  14) C  13                  0.00001     1.87078
 170. RY*(  15) C  13                  0.00000     1.76332
 171. RY*(  16) C  13                  0.00000     1.82263
 172. RY*(  17) C  13                  0.00000     1.50811
 173. RY*(   1) H  14                  0.00066     0.72233
 174. RY*(   2) H  14                  0.00016     1.51114
 175. RY*(   3) H  14                  0.00009     1.75902
 176. RY*(   4) H  14                  0.00008     1.88364
 177. RY*(   5) H  14                  0.00003     2.39098
 178. RY*(   1) H  15                  0.00040     0.89751
 179. RY*(   2) H  15                  0.00014     1.51213
 180. RY*(   3) H  15                  0.00008     1.60477
 181. RY*(   4) H  15                  0.00007     1.93871
 182. RY*(   5) H  15                  0.00002     2.33120
 183. RY*(   1) H  16                  0.00143     0.83676
```

184.	RY*(	2)	H	16	0.00016	1.58316
185.	RY*(	3)	H	16	0.00012	1.77873
186.	RY*(	4)	H	16	0.00008	1.67754
187.	RY*(	5)	H	16	0.00004	2.45028
188.	RY*(	1)	N	17	0.00285	1.08946
189.	RY*(	2)	N	17	0.00169	0.65025
190.	RY*(	3)	N	17	0.00089	0.92015
191.	RY*(	4)	N	17	0.00055	1.65709
192.	RY*(	5)	N	17	0.00046	0.98720
193.	RY*(	6)	N	17	0.00023	1.43044
194.	RY*(	7)	N	17	0.00020	1.52044
195.	RY*(	8)	N	17	0.00007	0.80134
196.	RY*(	9)	N	17	0.00005	1.24838
197.	RY*(	10)	N	17	0.00004	1.47774
198.	RY*(	11)	N	17	0.00004	3.92572
199.	RY*(	12)	N	17	0.00003	2.28934
200.	RY*(	13)	N	17	0.00001	2.36260
201.	RY*(	14)	N	17	0.00001	1.82682
202.	RY*(	15)	N	17	0.00000	34.63686
203.	RY*(	16)	N	17	0.00000	3.62227
204.	RY*(	17)	N	17	0.00001	3.87284
205.	RY*(	1)	H	18	0.00057	1.80570
206.	RY*(	2)	H	18	0.00044	1.39874
207.	RY*(	3)	H	18	0.00027	1.42243
208.	RY*(	4)	H	18	0.00014	1.77812
209.	RY*(	5)	H	18	0.00003	2.04452
210.	RY*(	1)	C	19	0.00631	1.32279
211.	RY*(	2)	C	19	0.00248	0.74032
212.	RY*(	3)	C	19	0.00213	0.80689
213.	RY*(	4)	C	19	0.00107	1.13211
214.	RY*(	5)	C	19	0.00078	1.36628
215.	RY*(	6)	C	19	0.00061	0.97774
216.	RY*(	7)	C	19	0.00045	1.38278
217.	RY*(	8)	C	19	0.00033	1.07312
218.	RY*(	9)	C	19	0.00015	2.18177
219.	RY*(	10)	C	19	0.00013	2.16968
220.	RY*(	11)	C	19	0.00011	1.00705
221.	RY*(	12)	C	19	0.00008	1.72065
222.	RY*(	13)	C	19	0.00005	3.49258
223.	RY*(	14)	C	19	0.00004	1.84636
224.	RY*(	15)	C	19	0.00001	1.85806
225.	RY*(	16)	C	19	0.00001	19.46453
226.	RY*(	17)	C	19	0.00001	2.51779
227.	RY*(	1)	C	20	0.00451	1.11885
228.	RY*(	2)	C	20	0.00400	0.59792
229.	RY*(	3)	C	20	0.00244	1.18777
230.	RY*(	4)	C	20	0.00225	1.06325
231.	RY*(	5)	C	20	0.00138	0.46156
232.	RY*(	6)	C	20	0.00075	0.45238
233.	RY*(	7)	C	20	0.00026	2.17159
234.	RY*(	8)	C	20	0.00021	1.49572
235.	RY*(	9)	C	20	0.00010	2.68676
236.	RY*(	10)	C	20	0.00008	1.94717
237.	RY*(	11)	C	20	0.00004	1.14072
238.	RY*(	12)	C	20	0.00003	2.27389
239.	RY*(	13)	C	20	0.00001	1.33505
240.	RY*(	14)	C	20	0.00000	1.76580
241.	RY*(	15)	C	20	0.00000	2.25587

242.	RY*(	16)	C	20					0.00000	20.36941	
243.	RY*(	17)	C	20					0.00001	0.79352	
244.	RY*(	1)	H	21					0.00037	1.20538	
245.	RY*(	2)	H	21					0.00028	1.43341	
246.	RY*(	3)	H	21					0.00006	1.46838	
247.	RY*(	4)	H	21					0.00005	1.62042	
248.	RY*(	5)	H	21					0.00003	2.27347	
249.	RY*(	1)	N	22					0.00264	1.10422	
250.	RY*(	2)	N	22					0.00145	0.66169	
251.	RY*(	3)	N	22					0.00076	1.62869	
252.	RY*(	4)	N	22					0.00070	0.95964	
253.	RY*(	5)	N	22					0.00042	1.08499	
254.	RY*(	6)	N	22					0.00024	1.55608	
255.	RY*(	7)	N	22					0.00015	1.55969	
256.	RY*(	8)	N	22					0.00004	3.57232	
257.	RY*(	9)	N	22					0.00005	1.55214	
258.	RY*(	10)	N	22					0.00002	1.39949	
259.	RY*(	11)	N	22					0.00002	2.05119	
260.	RY*(	12)	N	22					0.00002	1.05826	
261.	RY*(	13)	N	22					0.00000	2.53447	
262.	RY*(	14)	N	22					0.00000	1.86616	
263.	RY*(	15)	N	22					0.00000	34.74458	
264.	RY*(	16)	N	22					0.00000	3.33560	
265.	RY*(	17)	N	22					0.00000	3.29363	
266.	RY*(	1)	C	23					0.00422	1.15721	
267.	RY*(	2)	C	23					0.00229	0.64689	
268.	RY*(	3)	C	23					0.00195	0.80618	
269.	RY*(	4)	C	23					0.00072	0.72550	
270.	RY*(	5)	C	23					0.00053	0.60632	
271.	RY*(	6)	C	23					0.00034	0.81411	
272.	RY*(	7)	C	23					0.00023	0.79153	
273.	RY*(	8)	C	23					0.00016	1.67711	
274.	RY*(	9)	C	23					0.00012	2.52390	
275.	RY*(	10)	C	23					0.00007	1.84152	
276.	RY*(	11)	C	23					0.00006	2.34431	
277.	RY*(	12)	C	23					0.00003	1.71967	
278.	RY*(	13)	C	23					0.00004	1.68100	
279.	RY*(	14)	C	23					0.00001	1.44368	
280.	RY*(	15)	C	23					0.00000	19.95661	
281.	RY*(	16)	C	23					0.00001	2.75559	
282.	RY*(	17)	C	23					0.00001	2.62740	
283.	RY*(	1)	H	24					0.00033	1.31745	
284.	RY*(	2)	H	24					0.00020	1.46015	
285.	RY*(	3)	H	24					0.00018	1.63467	
286.	RY*(	4)	H	24					0.00005	1.48714	
287.	RY*(	5)	H	24					0.00003	2.25730	
304.	BD*(	1)	H	12	-	N	22		0.01272	0.20727	
305.	BD*(	1)	C	13	-	H	14		0.00789	0.24526	
306.	BD*(	1)	C	13	-	H	15		0.00263	0.24929	
307.	BD*(	1)	C	13	-	H	16		0.00933	0.24943	
308.	BD*(	1)	C	13	-	C	19		0.01091	0.26684	
309.	BD*(	1)	N	17	-	H	18		0.01427	0.20378	
310.	BD*(	1)	N	17	-	C	19		0.03014	0.26137	
311.	BD*(	1)	N	17	-	C	20		0.01235	0.28833	
312.	BD*(	1)	C	19	-	C	23		0.01817	0.40334	
313.	BD*(	2)	C	19	-	C	23		0.25415	-0.15158	

316(v),307(v),305(v),270(g)

217(g),271(g)

```
 314. BD*(   1) C  20 - H  21          0.00922     0.20272
 315. BD*(   1) C  20 - N  22          0.01269     0.28939
 316. BD*(   2) C  20 - N  22          0.52371    -0.23854
313(v),228(g),251(g)
 317. BD*(   1) N  22 - C  23          0.01164     0.26067
 318. BD*(   1) C  23 - H  24          0.00882     0.24163
       -------------------------------
             Total Lewis   43.00157  ( 97.7131%)
       Valence non-Lewis    0.93863  (  2.1329%)
       Rydberg non-Lewis    0.06777  (  0.1540%)
       -------------------------------
            Total unit  2   44.00797  (100.0000%)
           Charge unit  2    0.99203
Second Order Perturbation Theory Analysis of Fock Matrix in NBO Basis

     Threshold for printing:   0.50 kcal/mol
    (Intermolecular threshold: 0.05 kcal/mol)

E(2)  E(j)-E(i) F(i,j)
         Donor NBO (i)                    Acceptor NBO (j)
kcal/mol   a.u.    a.u.

================================================================================
============================

 within unit  1
   1. BD (   1) C   1 - H   2        / 73. RY*(   3) C   4
0.52    1.53    0.025
   1. BD (   1) C   1 - H   2        /297. BD*(   1) C   4 - H   5
2.86    0.88    0.045
   2. BD (   1) C   1 - H   3        /298. BD*(   1) C   4 - H   6
2.78    0.88    0.044
   3. BD (   1) C   1 - C   4        /298. BD*(   1) C   4 - H   6
0.54    1.01    0.021
   4. BD (   1) C   1 - H  25        /299. BD*(   1) C   4 - S   7
5.12    0.65    0.052
   5. BD (   1) C   4 - H   5        / 98. RY*(   1) S   7
0.88    1.04    0.027
   5. BD (   1) C   4 - H   5        /293. BD*(   1) C   1 - H   2
2.59    0.92    0.044
   6. BD (   1) C   4 - H   6        / 45. RY*(   2) C   1
0.62    1.13    0.024
   6. BD (   1) C   4 - H   6        / 99. RY*(   2) S   7
0.94    1.20    0.030
   6. BD (   1) C   4 - H   6        /294. BD*(   1) C   1 - H   3
2.74    0.92    0.045
   6. BD (   1) C   4 - H   6        /300. BD*(   1) S   7 - C   8
1.67    0.68    0.030
   7. BD (   1) C   4 - S   7        / 44. RY*(   1) C   1
0.92    1.26    0.030
   7. BD (   1) C   4 - S   7        /119. RY*(   1) C   8
0.94    1.42    0.033
   7. BD (   1) C   4 - S   7        /296. BD*(   1) C   1 - H  25
1.80    1.01    0.038
   7. BD (   1) C   4 - S   7        /303. BD*(   1) C   8 - H  11
1.24    0.98    0.031
   8. BD (   1) S   7 - C   8        / 71. RY*(   1) C   4
1.17    1.43    0.037
```

```
   8. BD (   1) S   7 - C   8        /298. BD*(   1) C   4 - H   6
1.23    1.00    0.031
  10. BD (   1) C   8 - H  10        / 98. RY*(   1) S   7
0.70    1.05    0.024
  11. BD (   1) C   8 - H  11        / 99. RY*(   2) S   7
0.55    1.21    0.023
  11. BD (   1) C   8 - H  11        /299. BD*(   1) C   4 - S   7
1.43    0.70    0.028
  27. CR (   1) C   1                / 61. RY*(   1) H   2
0.55   11.08    0.069
  27. CR (   1) C   1                / 72. RY*(   2) C   4
1.19   10.98    0.102
  28. CR (   1) C   4                / 47. RY*(   4) C   1
0.73   11.16    0.080
  28. CR (   1) C   4                /299. BD*(   1) C   4 - S   7
1.04   10.24    0.093
  30. CR (   2) S   7                /122. RY*(   4) C   8
0.51    9.27    0.062
  34. CR (   1) C   8                /137. RY*(   2) H   9
0.60   12.27    0.077
  34. CR (   1) C   8                /300. BD*(   1) S   7 - C   8
1.30   10.24    0.103
  41. LP (   1) S   7                /295. BD*(   1) C   1 - C   4
0.79    1.02    0.025
  41. LP (   1) S   7                /297. BD*(   1) C   4 - H   5
0.54    1.01    0.021
  41. LP (   1) S   7                /302. BD*(   1) C   8 - H  10
0.63    1.01    0.022
  42. LP (   2) S   7                / 72. RY*(   2) C   4
0.59    1.21    0.024
  42. LP (   2) S   7                /120. RY*(   2) C   8
0.79    1.63    0.033
  42. LP (   2) S   7                /295. BD*(   1) C   1 - C   4
4.06    0.63    0.046
  42. LP (   2) S   7                /296. BD*(   1) C   1 - H  25
0.83    0.65    0.021
  42. LP (   2) S   7                /297. BD*(   1) C   4 - H   5
3.48    0.63    0.042
  42. LP (   2) S   7                /301. BD*(   1) C   8 - H   9
5.10    0.63    0.051
  42. LP (   2) S   7                /302. BD*(   1) C   8 - H  10
3.72    0.63    0.044

 from unit  1 to unit  2
  41. LP (   1) S   7                /307. BD*(   1) C  13 - H  16
0.06    0.98    0.007
  41. LP (   1) S   7                /313. BD*(   2) C  19 - C  23
0.12    0.58    0.008
  41. LP (   1) S   7                /316. BD*(   2) C  20 - N  22
0.09    0.49    0.007
  42. LP (   2) S   7                /183. RY*(   1) H  16
0.11    1.18    0.010
  42. LP (   2) S   7                /185. RY*(   3) H  16
0.09    2.12    0.013
  42. LP (   2) S   7                /231. RY*(   5) C  20
0.05    0.81    0.006
  42. LP (   2) S   7                /250. RY*(   2) N  22
0.08    1.01    0.008
```

```
  42. LP (   2) S   7              /307. BD*(   1) C  13 - H  16
0.06    0.59    0.005
  42. LP (   2) S   7              /313. BD*(   2) C  19 - C  23
0.22    0.19    0.006
  42. LP (   2) S   7              /316. BD*(   2) C  20 - N  22
0.14    0.11    0.004

 from unit  2 to unit  1
  21. BD (   2) C  19 - C  23      /299. BD*(   1) C   4 - S   7
0.32    0.51    0.012
  43. LP (   1) N  17              /300. BD*(   1) S   7 - C   8
0.08    0.52    0.006

 within unit  2
  12. BD (   1) H  12 - N  22      /227. RY*(   1) C  20
1.03    2.00    0.041
  12. BD (   1) H  12 - N  22      /266. RY*(   1) C  23
1.01    2.04    0.041
  12. BD (   1) H  12 - N  22      /311. BD*(   1) N  17 - C  20
1.39    1.17    0.036
  12. BD (   1) H  12 - N  22      /312. BD*(   1) C  19 - C  23
1.19    1.29    0.035
  13. BD (   1) C  13 - H  14      /212. RY*(   3) C  19
0.64    1.46    0.028
  13. BD (   1) C  13 - H  14      /312. BD*(   1) C  19 - C  23
2.85    1.06    0.049
  13. BD (   1) C  13 - H  14      /313. BD*(   2) C  19 - C  23
3.99    0.50    0.042
  14. BD (   1) C  13 - H  15      /310. BD*(   1) N  17 - C  19
7.50    0.91    0.074
  15. BD (   1) C  13 - H  16      /312. BD*(   1) C  19 - C  23
2.34    1.06    0.044
  15. BD (   1) C  13 - H  16      /313. BD*(   2) C  19 - C  23
5.06    0.50    0.048
  16. BD (   1) C  13 - C  19      /188. RY*(   1) N  17
0.94    1.92    0.038
  16. BD (   1) C  13 - C  19      /266. RY*(   1) C  23
1.02    1.99    0.040
  16. BD (   1) C  13 - C  19      /310. BD*(   1) N  17 - C  19
0.81    1.09    0.027
  16. BD (   1) C  13 - C  19      /311. BD*(   1) N  17 - C  20
2.42    1.12    0.046
  16. BD (   1) C  13 - C  19      /312. BD*(   1) C  19 - C  23
3.16    1.23    0.056
  16. BD (   1) C  13 - C  19      /317. BD*(   1) N  22 - C  23
0.88    1.09    0.028
  17. BD (   1) N  17 - H  18      /210. RY*(   1) C  19
1.58    2.20    0.053
  17. BD (   1) N  17 - H  18      /227. RY*(   1) C  20
0.95    2.00    0.039
  17. BD (   1) N  17 - H  18      /310. BD*(   1) N  17 - C  19
0.55    1.14    0.022
  17. BD (   1) N  17 - H  18      /312. BD*(   1) C  19 - C  23
1.61    1.28    0.041
  17. BD (   1) N  17 - H  18      /315. BD*(   1) C  20 - N  22
1.40    1.17    0.036
  18. BD (   1) N  17 - C  19      /230. RY*(   4) C  20
1.37    2.07    0.048
```

```
  18. BD (   1) N  17 - C  19        /306. BD*(   1) C  13 - H  15
0.51    1.25    0.023
  18. BD (   1) N  17 - C  19        /308. BD*(   1) C  13 - C  19
0.77    1.27    0.028
  18. BD (   1) N  17 - C  19        /309. BD*(   1) N  17 - H  18
0.70    1.21    0.026
  18. BD (   1) N  17 - C  19        /311. BD*(   1) N  17 - C  20
1.69    1.29    0.042
  18. BD (   1) N  17 - C  19        /312. BD*(   1) C  19 - C  23
0.64    1.41    0.027
  18. BD (   1) N  17 - C  19        /314. BD*(   1) C  20 - H  21
2.51    1.21    0.049
  18. BD (   1) N  17 - C  19        /315. BD*(   1) C  20 - N  22
1.02    1.29    0.032
  18. BD (   1) N  17 - C  19        /317. BD*(   1) N  22 - C  23
0.84    1.26    0.029
  18. BD (   1) N  17 - C  19        /318. BD*(   1) C  23 - H  24
2.81    1.25    0.053
  19. BD (   1) N  17 - C  20        /210. RY*(   1) C  19
2.21    2.36    0.065
  19. BD (   1) N  17 - C  20        /211. RY*(   2) C  19
0.90    1.78    0.036
  19. BD (   1) N  17 - C  20        /212. RY*(   3) C  19
0.61    1.85    0.030
  19. BD (   1) N  17 - C  20        /304. BD*(   1) H  12 - N  22
2.65    1.25    0.051
  19. BD (   1) N  17 - C  20        /308. BD*(   1) C  13 - C  19
3.15    1.31    0.057
  19. BD (   1) N  17 - C  20        /309. BD*(   1) N  17 - H  18
0.58    1.25    0.024
  19. BD (   1) N  17 - C  20        /310. BD*(   1) N  17 - C  19
1.66    1.30    0.042
  19. BD (   1) N  17 - C  20        /314. BD*(   1) C  20 - H  21
0.51    1.24    0.022
  20. BD (   1) C  19 - C  23        /156. RY*(   1) C  13
0.83    1.41    0.031
  20. BD (   1) C  19 - C  23        /188. RY*(   1) N  17
0.60    1.99    0.031
  20. BD (   1) C  19 - C  23        /249. RY*(   1) N  22
0.63    2.01    0.032
  20. BD (   1) C  19 - C  23        /252. RY*(   4) N  22
0.62    1.86    0.031
  20. BD (   1) C  19 - C  23        /304. BD*(   1) H  12 - N  22
3.39    1.11    0.055
  20. BD (   1) C  19 - C  23        /308. BD*(   1) C  13 - C  19
2.65    1.17    0.050
  20. BD (   1) C  19 - C  23        /309. BD*(   1) N  17 - H  18
3.85    1.11    0.058
  20. BD (   1) C  19 - C  23        /317. BD*(   1) N  22 - C  23
0.62    1.16    0.024
  20. BD (   1) C  19 - C  23        /318. BD*(   1) C  23 - H  24
1.66    1.14    0.039
  21. BD (   2) C  19 - C  23        /157. RY*(   2) C  13
0.68    1.03    0.025
  21. BD (   2) C  19 - C  23        /189. RY*(   2) N  17
1.54    1.10    0.039
  21. BD (   2) C  19 - C  23        /250. RY*(   2) N  22
1.46    1.11    0.038
```

```
  21. BD (   2) C  19 - C  23        /305. BD*(   1) C  13 - H  14
2.33    0.69    0.038
  21. BD (   2) C  19 - C  23        /307. BD*(   1) C  13 - H  16
2.54    0.70    0.040
  21. BD (   2) C  19 - C  23        /316. BD*(   2) C  20 - N  22
17.16    0.21    0.058
  22. BD (   1) C  20 - H  21        /188. RY*(   1) N  17
0.65    1.86    0.031
  22. BD (   1) C  20 - H  21        /249. RY*(   1) N  22
0.72    1.87    0.033
  22. BD (   1) C  20 - H  21        /310. BD*(   1) N  17 - C  19
2.72    1.03    0.047
  22. BD (   1) C  20 - H  21        /317. BD*(   1) N  22 - C  23
2.60    1.03    0.046
  23. BD (   1) C  20 - N  22        /190. RY*(   3) N  17
0.53    1.97    0.029
  23. BD (   1) C  20 - N  22        /266. RY*(   1) C  23
1.54    2.20    0.052
  23. BD (   1) C  20 - N  22        /268. RY*(   3) C  23
1.05    1.85    0.039
  23. BD (   1) C  20 - N  22        /304. BD*(   1) H  12 - N  22
0.59    1.25    0.024
  23. BD (   1) C  20 - N  22        /309. BD*(   1) N  17 - H  18
2.63    1.25    0.051
  23. BD (   1) C  20 - N  22        /314. BD*(   1) C  20 - H  21
0.52    1.25    0.023
  23. BD (   1) C  20 - N  22        /317. BD*(   1) N  22 - C  23
1.33    1.31    0.037
  23. BD (   1) C  20 - N  22        /318. BD*(   1) C  23 - H  24
1.92    1.29    0.044
  24. BD (   2) C  20 - N  22        /267. RY*(   2) C  23
1.95    1.20    0.044
  24. BD (   2) C  20 - N  22        /313. BD*(   2) C  19 - C  23
15.36    0.40    0.073
  25. BD (   1) N  22 - C  23        /211. RY*(   2) C  19
0.51    1.74    0.027
  25. BD (   1) N  22 - C  23        /212. RY*(   3) C  19
0.57    1.81    0.029
  25. BD (   1) N  22 - C  23        /230. RY*(   4) C  20
1.38    2.07    0.048
  25. BD (   1) N  22 - C  23        /304. BD*(   1) H  12 - N  22
0.59    1.21    0.024
  25. BD (   1) N  22 - C  23        /308. BD*(   1) C  13 - C  19
4.23    1.27    0.065
  25. BD (   1) N  22 - C  23        /310. BD*(   1) N  17 - C  19
0.86    1.26    0.030
  25. BD (   1) N  22 - C  23        /311. BD*(   1) N  17 - C  20
1.17    1.29    0.035
  25. BD (   1) N  22 - C  23        /312. BD*(   1) C  19 - C  23
0.82    1.41    0.030
  25. BD (   1) N  22 - C  23        /314. BD*(   1) C  20 - H  21
2.70    1.20    0.051
  25. BD (   1) N  22 - C  23        /315. BD*(   1) C  20 - N  22
1.50    1.29    0.039
  26. BD (   1) C  23 - H  24        /210. RY*(   1) C  19
1.71    2.05    0.053
  26. BD (   1) C  23 - H  24        /249. RY*(   1) N  22
0.70    1.83    0.032
```

```
  26. BD (   1) C  23 - H  24        /310. BD*(   1) N  17 - C  19
2.05    0.99    0.040
  26. BD (   1) C  23 - H  24        /312. BD*(   1) C  19 - C  23
2.15    1.13    0.044
  26. BD (   1) C  23 - H  24        /315. BD*(   1) C  20 - N  22
2.88    1.02    0.048
  35. CR (   1) C  13                /178. RY*(   1) H  15
0.51   11.08    0.067
  35. CR (   1) C  13                /183. RY*(   1) H  16
0.60   11.02    0.072
  35. CR (   1) C  13                /210. RY*(   1) C  19
0.99   11.51    0.095
  35. CR (   1) C  13                /310. BD*(   1) N  17 - C  19
0.51   10.45    0.066
  35. CR (   1) C  13                /312. BD*(   1) C  19 - C  23
0.85   10.59    0.085
  36. CR (   1) N  17                /205. RY*(   1) H  18
0.85   16.20    0.105
  36. CR (   1) N  17                /211. RY*(   2) C  19
1.52   15.13    0.135
  36. CR (   1) N  17                /212. RY*(   3) C  19
1.32   15.20    0.127
  36. CR (   1) N  17                /230. RY*(   4) C  20
1.30   15.45    0.127
  37. CR (   1) C  19                /158. RY*(   3) C  13
0.84   11.20    0.087
  37. CR (   1) C  19                /190. RY*(   3) N  17
0.53   11.19    0.068
  37. CR (   1) C  19                /266. RY*(   1) C  23
0.63   11.42    0.076
  37. CR (   1) C  19                /268. RY*(   3) C  23
0.62   11.07    0.074
  37. CR (   1) C  19                /311. BD*(   1) N  17 - C  20
0.76   10.56    0.080
  37. CR (   1) C  19                /312. BD*(   1) C  19 - C  23
0.77   10.67    0.081
  37. CR (   1) C  19                /318. BD*(   1) C  23 - H  24
0.68   10.51    0.076
  38. CR (   1) C  20                /244. RY*(   1) H  21
0.70   11.53    0.080
  38. CR (   1) C  20                /252. RY*(   4) N  22
0.78   11.28    0.084
  38. CR (   1) C  20                /310. BD*(   1) N  17 - C  19
0.96   10.58    0.091
  38. CR (   1) C  20                /317. BD*(   1) N  22 - C  23
0.85   10.58    0.085
  39. CR (   1) N  22                /151. RY*(   1) H  12
0.72   16.71    0.098
  39. CR (   1) N  22                /230. RY*(   4) C  20
1.32   15.46    0.128
  39. CR (   1) N  22                /268. RY*(   3) C  23
1.74   15.20    0.145
  40. CR (   1) C  23                /210. RY*(   1) C  19
1.55   11.58    0.120
  40. CR (   1) C  23                /211. RY*(   2) C  19
0.58   10.99    0.071
  40. CR (   1) C  23                /283. RY*(   1) H  24
0.75   11.57    0.083
```

```
  40. CR (   1) C  23              /308. BD*(   1) C  13 - C  19
1.35   10.52    0.107
  40. CR (   1) C  23              /312. BD*(   1) C  19 - C  23
0.85   10.66    0.085
  40. CR (   1) C  23              /315. BD*(   1) C  20 - N  22
0.75   10.54    0.080
  43. LP (   1) N  17              /189. RY*(   2) N  17
0.66    1.11    0.028
  43. LP (   1) N  17              /206. RY*(   2) H  18
0.80    1.86    0.040
  43. LP (   1) N  17              /211. RY*(   2) C  19
0.97    1.20    0.035
  43. LP (   1) N  17              /212. RY*(   3) C  19
0.59    1.26    0.028
  43. LP (   1) N  17              /214. RY*(   5) C  19
0.71    1.82    0.037
  43. LP (   1) N  17              /229. RY*(   3) C  20
2.22    1.65    0.062
  43. LP (   1) N  17              /313. BD*(   2) C  19 - C  23
32.09    0.31    0.094
  43. LP (   1) N  17              /316. BD*(   2) C  20 - N  22
77.50    0.22    0.118
 313. BD*(   2) C  19 - C  23      /217. RY*(   8) C  19
0.60    1.22    0.068
 313. BD*(   2) C  19 - C  23      /270. RY*(   5) C  23
0.67    0.76    0.056
 313. BD*(   2) C  19 - C  23      /271. RY*(   6) C  23
0.54    0.97    0.057
 313. BD*(   2) C  19 - C  23      /305. BD*(   1) C  13 - H  14
0.83    0.40    0.045
 313. BD*(   2) C  19 - C  23      /307. BD*(   1) C  13 - H  16
0.94    0.40    0.048
 316. BD*(   2) C  20 - N  22      /228. RY*(   2) C  20
4.63    0.84    0.108
 316. BD*(   2) C  20 - N  22      /251. RY*(   3) N  22
0.77    1.87    0.066
 316. BD*(   2) C  20 - N  22      /313. BD*(   2) C  19 - C  23
12.94    0.09    0.048
```

8. His-Met structural searches

8.1 Protein Data Bank

The co-ordinates of the interacting histidine and methionine from the QM/MM calculations were taken as a seed for a JESS[13] template search of a set of 2540 structures culled from the Protein Data Bank (http://www.pdb.org)[14] using PISCES[15]. Each structure had resolution of ≤ 1.6 Å and shared no more than 30% sequence identity with any other structure in the set. Briefly, JESS uses a *kd*-tree data structure and a back-tracking constraint solving algorithm to search structures for constellations of atoms, within a defined RMSD cut-off, which for this search was set to 1.0 Å. Only identical atom matches were allowed. Figure 3f in the main article shows the top ten hits resulting from the search.

8.2 Cambridge Structural Database (CSD)

8.2.1 Initial CSD search for ab initio optimized His-Met pair

The Cambridge Structural Database [CSD version 5.33 (November 2011) including three updates (May2012), using ConQuest version 1.14] contains 76 structures with a 4-methylimidazole fragment (substituents on N and C unspecified, with 9 also containing a sulfur atom ≤ 5 Å (intermolecular) away from the methylimidazole/imidazolium ring centroid [CSD reference codes CAQYEY, CUPNEH, CUPNIL, CUPNOR, GEZYAL, GEZYEP, POMBEY, SOXMUN and YESHAF]. Manual inspection of these crystal structures revealed that none contain the imidazolium···C-S-C motif computed for the Met-HisH$^+$ model (Figure 3). However, five entries did contain related features (Figure S3). For example, the structures CUPNEH, CUPNIL and CUPNOR contain a distorted dimethyl sulfoxide (DMSO) molecule that is located very near an imidazolium ring. The shortest distance found is between the DMSO O-atom and one of the imidazolium ring atoms and the atoms are at a separation less than the sum of their van der Waals radii ($\Sigma_{vdWaals(O+C)}$ = 3.22 Å), strongly suggestive of a bonding interaction. These supramolecular features found in CUPNEH, CUPNIL, CUPNOR, POMBEY and YESHAF have not been commented on in the original publications[16-18].

8.2.2 Directionality of face-on interactions involving (neutral, protonated or coordinated) imidazole

A model of the imidazole(ium) fragment was developed with geometric parameters as shown in Figure S4a. The distances m, n, o, p, q and r, as well as the angles *α*, *β*, *γ* and *ε* were measured using ConQuest; their average values and standard deviations are listed in Table S4. Based on these measurements, assuming that [*x*, *y*] in C^1 is [0, 0], and that C^1 – H^1 is on the *y*-axis, the *x* and *y* values for all atoms can be derived as follows (see Figure S4a-c, Table S5).

The pentagon C^1-N^1-C^2-C^3-N^2 is confined within a square with $x = 2m^x$ and $y = m^y+n^y$, meaning that the *x*-value for N^1 and N^2 is $\pm m^x$, and that the *y*-value for C^2 and C^3 is $-(m^y+n^y)$. The *y*-value for N^1 and N^2 is $-m^y$, while the *x*-value for C^2 and C^3 is $\pm(0.5o+n^x)$. As *α* and *β* are known, *α*' and *β*' are know as well, with which m^x, m^y, n^x and n^y can thus be computed. H^1 is naturally located on [0, r]. The $C^{2/3}$ – $H^{2/3}$ distance p can be seen as the diagonal through a rectangle with $x = p^x$ and $y = p^y$. Likewise, the $N^{1/2}$ – $X^{1/2}$ distance q can be seen as the diagonal of rectangle with $x = q^x$ and $y = q^y$. Since both *γ* and *ε* are known, *γ'* and *ε'* are known as well, from which p^x, p^y, q^x and q^y can thus be derived. The *x*-value for $H^{2/3}$ then is

$\pm(m^x+p^x)$; its *y*-value is $-(m^y+n^y+p^y)$. $X^{1/2}$ is located on $x = \pm(m^x+q^x)$ and $y = -m^x+q^y$. Finally, the ring centroid can be derived by taking the average *x*- and *y*-values of C^1-N^1-C^2-C^3-N^2, which numerically results in [0, -1.125]. For the construction of our model, it was assumed that all atoms in the ring, as well as the ring centroid, were perfectly aligned with the *xy* plane, so that in all cases $z = 0$. The *x*, *y*, *z*-coordinates of all atoms belonging to the model ring are thus known and given in Table S5.

To generate the space-filling model used in our analysis, the van der Waals (vdW) radii of the atoms must be known as well. The vdW radii of C (1.70 Å), N (1.55 Å) and H (1.09 Å) are known[19]. Because the identity of $X^{1/2}$ was unspecified, its identity was measured using ConQuest, wherefrom the average vdW radius was computed (see Table S5) by assuming van vdW radii listed elsewhere[19]. Using these vdW radii, the individual atoms were drawn as parts (.ipt extension) using Autodesk® Inventor® Professional 2012. The parts were then placed into an assembly file (.iam extension), where they were positioned so that the ring centroid and the [x, y, z] centre of the assembly file were aligned. The parts in the assembly file were merged into one body using the 'Shrinkwrap' function in Autodesk® Inventor®, resulting in a new part file (.ipt extension); this is thus the model used (Figure S4d-e).

The Cartesian coordinates of the electron rich atom (El.R.) that is potentially interacting with an imidazole(ium) ring were initially determined relative to C^1 as point [0,0,0] (See Figure S4b-d). For this purpose, two parallel displacement parameters were computed (Figure S4b-c) using Pythagoras' theorem; *r* (relative to the ring cetroid) was computed based on the ring centroid – El.R. distance (*D*) and the ring plane – El.R. distance (*d*, see also Figure S4f); *r'* was derived from the C^1 – El.R. distance and the ring plane – El.R. distance. These two distances (*r* and *r'*) together with the distance from C^1 to the ring centroid (which is known), thus make up the irregular triangle (T1) represented by the red dashed lines in Figure S4c. As the length of all three sides (labelled *r*, *r'* and *c*) of this triangle are known, its angles can be calculated. The angle relevant for present purposes, namely *α*, can thus be derived, from which *α'* follows. This angle (*α'*) together with the steep side (*r'*) of the regular triangle (T2) shown in grey in Figure S4c then enable the computation of *x* and *y* according to equations S1 and S2 respectively:

Eqn. S1. $$\sin\left(180-\arccos\left(\frac{(r')^2+c^2+r^2}{2r'c}\right)\right)r'$$

Eqn. S2. $$\cos\left(180-\arccos\left(\frac{(r')^2+c^2+r^2}{2r'c}\right)\right)r'$$

When the electron rich atom was situated directly above the *y*-axis, *α* could not be computed. In these instances, when $r' = r$, then $y = -r' = -r$; when $r' < r$ and $r > 1.125$, then $y = r'$; in all other cases $y = -r'$. The *z*-axis is given by the distance from the electron rich atom to the ring plane. Adding 1.125 to the *y*-values of the data thus computed gives the *y*-value relative to the ring centroid.

An electron rich atom ('El.R.'; halogen, N, O, P, S, As, Se, Te) was considered as potentially interacting with our model ring when the *inter*molecular distance D was ≤ 5 Å relative to the ring centre of our model (Figure S4f). Other measurements used in our statistical analysis are the C^1–ring centroid–El.R. angle (α) and the El.R.–ring plane distance (d). Using Pythagoras' theorem, the parallel displacement parameter (r) could be derived from D and d. To inquire after the directional nature of the interaction between an electron rich atom (generally) and our model imidazole(ium) ring, two separate statistical analyses were performed. The methodologies used for these analyses are illustrated in Figure S4g-j and Figure S4k-m, and concern the evaluation of CSD data in the *xy*- and *yz*-plane of the imidazole(ium) model respectively.

As D (see Figure S4f) can be 5 Å maximally, the hits retrieved from the CSD are contained within a sphere with a radius of 5 Å. The subset of these data wherein $d \leq 2$ Å is thus contained within a spherical segment with a width of 10 Å and a height of 4 Å (2 Å above and below the *xy* plane, see Figure S4g-j). This spherical segment can be divided into geometric objects as a function of α. The data below a certain value of α are thus located within a spherical cone with cone angle α. As an example, Figure S4g-i shows such a spherical cone with a cone angle of 5º. At certain values of α however, the base of the spherical cone will exceed the boundary set by the spherical segment. Thus, the body in which the data can be located equals a spherical cone that has the 'top' (+z) and 'bottom' (–z) sliced off (in between $z = \pm$ 2Å) so that the body fits the spherical segment; Figure S4j shows an example where α = 60º.

The 'free' volume in between two values of α wherein an electron rich atom can *actually* be located, is given by the volume of the body in between those two values of α, minus the volume that the model imidazole(ium) occupies there. To derive these data, the bodies in between α_a and α_b were drawn in Autodesk® Inventor® Professional, with 5º increments, until α = 90º. Mirroring these bodies along the *xz* plane then gives the bodies for α = 90–180º. The volume of the model ring as a function of α was also computed using Autodesk® Inventor® Professional, by calculating the interfering volume between the model ring and the bodies with α = 0–180º (with 5º increments).

The volume fraction ($V_{a\text{-}b}$) could thus be computed by dividing the 'free' volume in between $\alpha_{a\text{-}b}$ by the total 'free' volume within the 10 Å wide and 4 Å high spherical segment. Likewise, the hit fraction ($H_{a\text{-}b}$) can be obtained by dividing the number of hits with $\alpha_{a\text{-}b}$ by the total number of hits found with $D \leq 5$ and $r \leq 2$ Å. The ratio $H_{a\text{-}b}/V_{a\text{-}b}$ represents the probability ($P_{a\text{-}b}$) to find a hit within each volume considered (in between $\alpha_{a\text{-}b}$). By default, P equals unity for a random distribution; for attractive interactions, P is expected to be > 1, while P < 1 is indicative of a repulsive interaction. Plotting P as a function of α will thus reveal any *non-accidental* special distribution of hits around the periphery of our model in between α = 0º (on the y-axis, next to H^1) and 180º (on the y-axis, in between H^2 and H^3), where the data at α = 90º are located in the *xz* plane (see Figure S4g and h). Due to the symmetric nature of the model, such plots reveal the data for the asymmetric unit only (so in between $|z| \leq 2$ Å and $|x| \leq 5$ Å).

The method used for our second analysis is similar to – but distinctly different from – the one described above. The subset of the data originally retrieved from the CSD ($D \leq 5$ Å) wherein $|x| \leq 2$ Å) contains hits that are located within a 10 Å wide and 4 Å high spherical segment that is centered at the *zy*-plane (see Figure S4k-m). This spherical segment can be

divided into geometric objects as a function of *r*. The data below a certain value of *r* are located within a body that is a cylinder plus two spherical caps (top and bottom), all with a basal radius of *r*. As an example, Figure S4k-m shows such a body with *r* = 0.25 Å. Above a certain value of *r* however, the edges of the body will exceed the boundary set by the spherical segment. Thus, the body in which the data can be located equals a cylinder and two spherical caps that has the 'left' (–*x*) and 'right' (+*x*) sides sliced off (in between *x* = ± 2 Å) so that the body fits the spherical segment. The 'free' volume in between two values of *r* wherein an electron rich atom can *actually* be located, is given by the volume of the body in between those two values of *r*, minus the volume that the imidazole(ium) model occupies there. Due to the asymmetric nature of the model along the *zy* plane, the bodies that were actually drawn in Autodesk® Inventor® Professional were a half cylinder with the corresponding half spherical caps on the top and bottom. Such bodies were generated for *r* = 0 – 5 with 0.25 Å increments, and centred at the *x*-axis with the round side pointing in between the two N-atoms (so towards H^1); the data that is found within these bodies is thus also characterized by $\alpha < 90^o$ (see Figure S4l). Mirroring these bodies along the *xz*-plane then gives similar bodies wherein the round sides are pointing away from the two N-atoms (so in between H^2 and H^3); the data that is found within these bodies is thus also characterized by $\alpha > 90^o$ (see Figure S4l). The volumes of these (10/0.25=) 40 bodies, as well as the volume of the model ring within these bodies was computed as a function of *r* (and *α*) using Autodesk® Inventor® Professional.

The volume fraction ($V_{a\text{-}b}$) could thus be determined. After determining the hit fraction ($H_{a\text{-}b}$) of the CSD data in between $|x| \leq 2$ Å and $|z| \leq 5$ Å, the Probability ($P_{a\text{-}b}$) could be computed as a function of *r*. Again, *P* should by default be unity for a random distribution, while $P > 1$ and $P < 1$ are indicative of attractive and repulsive interactions respectively. Plotting *P* as a function of *r* will thus reveal any *non-accidental* special distribution of hits around our model in between $|r| = 5$ Å; $\alpha = 0^o$ (on the y-axis, next to H^1) and $|r| = 5$ Å; $\alpha = 180^o$ (on the y-axis, in between H^2 and H^3). The data at $r \leq 2.5$ Å are located above/below the ring in the *xy* plane (see Figure S4l). Due to the symmetric nature of the model, such plots reveal the data for the asymmetric unit only (so for $|x| \leq 2$ Å and $|z| \leq 5$ Å).

The *P vs*. *α* or *r* plots for these two distinct analyses are shown respectively in Figure S5a and S5b. Figure S5c-f shows the *P vs*. *r* plots of more specific datasets (i.e. where El.R. is specified). Figure S5h-i sows plots of the hit fraction (within bodies selected based on Figure S4k-m) as a function of *D* (Figure S5h) or *d* (Figure S5i), which can reveal any potential overlap of van der Waals shells of interacting partners (highly suggestive of bonding interactions).

These data suggest that the most common mode of interaction between an electron rich atom and an imidazole(ium) ring is through hydrogen bonding ($P > 1$ @ $\alpha < 30°$ and $> 150°$ (Figure S5a) and @ $r > 4$ Å (Figure S5b)). Generally (N as El.R. is the exception, see Figure S5g) H1 is a better hydrogen bond donor than H2 and H3 (i.e. the probability *P* of a hit being within a certain volume at a displacement $r = 4 – 5$ Å is higher at $\alpha < 90^o$ than at $\alpha > 90^o$ and more hits fall within the benchmark for van der Waals overlap (Figure S5h)). Although less pronounced than H-bonding, the El.R. – *π* interaction is evident as a general motif as well, as $P > 1$ around $r = 0$ (see Figure S5b-d). Actually, examination of the hit fraction within the region defined by $r \leq 0.5$ Å revealed that about 20% (integration was carried out till the nearest data point of 3.2 Å) of the El.R atoms are within a distance highly indicative of van der Waals overlap with the central imidazole(ium) ring (i.e. C + F = 3.17 Å, see Figure S5i).

The structures that contain the molecule pairs complying to these strict criteria (i.e. $r \leq 0.5$ Å and $d \leq 3.2$ Å) are: HICDAZ (d = 3.051 Å), JIMWAE (d = 3.147 Å), NODKUN (d = 3.002 Å), GEWWEX (d = 3.055; 2.897 Å), HIPYOV (d = 2.859 Å), JETFEU (d = 2.886 Å), JETFIY (d = 3.071; 3.018 Å), NAQMEY (d = 3.164; 3.142 Å), NAQMOI (d = 2.938; 2.981 Å), REXCUT (d

Table S1. Average interatomic distances for hydrogen bonds identified in the active site of PTDH during two separate 1 ns AM1/CHARMM27 MD simulations. Distances (in Å) are the average over 2000 snapshots and SD is one standard deviation of the distance.

	Trajectory 1		**Trajectory 2**	
	Distance	**SD**	**Distance**	**SD**
His292 HD1-GLU266 OE1	1.92	0.23	2.03	0.38
His292 HD1-GLU266 OE2	2.87	0.30	2.03	0.38
Phosphite O1-Arg237 HH12	2.26	0.68	1.83	0.12
Phosphite O1-Arg237 HH22	2.13	0.46	2.01	0.18
Phosphite O1-Lys76 HN	3.34	0.69	4.22	0.15
Phosphite O1-Gly77 HN	2.42	0.51	3.17	0.21
Phosphite O2-Arg237 HH22	2.67	0.95	2.73	0.75
Phosphite O2-Arg237 HH22	4.19	0.27	3.95	0.22
Phosphite O2-Lys76 HN	2.21	0.64	1.89	0.12
Phosphite O2-Gly77 HN	3.24	0.86	1.92	0.15
Phosphite H2-Met53 SD	3.24	0.93	2.52	0.26
Arg237 HE-Asp79 OD2	2.03	0.26	1.89	0.21
Arg237 HH21-Asp79 OD1	1.87	0.22	1.82	0.16
Arg237 HH21-Asp79 OD2	3.21	0.36	2.98	0.32

Table S2. Small molecule calculations with various QM methods to assess the accuracy of the AM1 method for the calculation of the reaction energy for the nucleophilic displacement reaction catalyzed by PTDH (Figure 1).

	AM1	**PM3**	**B3LYP/ 6-31+G(d)**	**MP2/ 6-31+G(d)**
Protonated phosphite	-0.363	-0.340	-568.369	-567.268
4-methyl imidazole	0.067	0.034	-265.548	-264.709
Nicotinamide ring of NAD+	0.241	0.242	-417.365	-416.110
Water	-0.094	-0.085	-76.423	-76.210
Di-protonated phosphate	-0.503	-0.464	-643.631	-642.356
4-methyl imidazolium	0.296	0.265	-265.924	-265.080
Nicotinamide ring of NADH	-0.020	-0.023	-418.185	-416.900
$E_{Reactants}$ **(H)**	-0.149	-0.149	-1327.705	-1324.297
$E_{Products}$ **(H)**	-0.227	-0.222	-1327.741	-1324.336
$\Delta_r E$ **(kcal/mol)**	-49.2	-45.5	-22.5	-24.3

Table S3. The average contributions of different active-site residues to (a) the (QM/MM) electrostatic and (b) the (QM/MM) van der Waals interaction energy along the reaction path calculated at the AM1/CHARMM27 level of theory. All energies are in kcal/mol and the numbers in bold indicate the residue contribution relative to the value in the substrate complex (R). Lys76 NH indicates the electrostatic energy of the interaction when only the charges of the backbone amide are set to zero (i.e. the +1*e* charge of the sidechain is retained). The interaction energy of Met53 is decomposed further by switching different electrostatic groups in the CHARMM27 representation of methionine to zero. Met53 CO indicates that the charges of the group containing the backbone carbonyl are set to zero. Met53 NH indicates that the group containing the backbone amide are set to zero and Met53 S indicates that the charges of the sidechain are set to zero.

(a)	R		TS1		Int		TS2		P	
TOTAL	−301.7	**0**	−305.5	**5.1**	−359.1	**−48.5**	−352.3	**−41.7**	−375.4	**−64.8**
Arg237	−49.5	**0**	−51.0	**−1.5**	−75.7	**−26.2**	−58.3	**−8.8**	−63.1	**−13.6**
Arg301	−19.5	**0**	−25.7	**−6.2**	−28.5	**−9.1**	−30.5	**−11.0**	−27.8	**−8.4**
Asp79	16.3	**0**	15.5	**−0.7**	21.8	**5.5**	16.2	**−0.1**	20.4	**4.1**
Glu266	−20.9	**0**	−26.4	**−5.6**	−49.9	**−29.1**	−53.8	**−33.0**	−59.8	**−38.9**
Gly77	−6.2	**0**	−3.8	**2.4**	−6.5	**−0.3**	−5.79	**0.4**	−8.0	**−1.8**
Lys76	−10.4	**0**	−14.7	**−4.4**	−29.5	**−19.1**	−34.1	**−23.7**	−39.1	**−28.7**
Lys76 NH	−6.5	**0**	−7.5	**−1.2**	−10.7	**−4.5**	−7.0	**−0.7**	−6.8	**−0.5**
Met53	1.0	**0**	2.2	**1.2**	1.9	**0.9**	−6.9	**−7.9**	−1.0	**−2.0**
Met53 CO	−0.5	**0**	−0.7	**−0.2**	−0.5	**−0.1**	−1.07	**−0.62**	0.08	**0.54**
Met53 NH	−0.8	**0**	−1.1	**−0.3**	−1.0	**−0.2**	−0.58	**0.25**	−0.94	**−0.11**
Met53 S	2.5	**0**	4.4	**1.9**	4.1	**1.6**	−5.03	**−7.54**	0.32	**−2.18**
Ser295	−1.7	**0**	−3.0	**−1.3**	−2.7	**−1.0**	−4.9	**−3.2**	−6.8	**−5.1**

(b)	R		TS1		Int		TS2		P	
TOTAL	−11.5	**0**	−13.6	**−2.1**	−7.7	**3.8**	−10.70	**0.81**	−15.73	**−4.22**
Arg237	−0.8	**0**	−0.3	**0.5**	0.0	**0.8**	−1.57	**−0.76**	−0.45	**0.35**
Arg301	−0.3	**0**	−0.4	**−0.1**	−0.4	**−0.1**	−0.38	**−0.07**	−0.33	**−0.02**
Asp79	−0.2	**0**	−0.1	**0.0**	−0.1	**0.0**	−0.13	**0.04**	−0.16	**0.00**
Glu266	−0.1	**0**	−0.2	**−0.2**	1.0	**1.1**	−0.26	**−0.18**	1.10	**1.17**
Gly77	−0.4	**0**	−1.5	**−1.0**	−1.3	**−0.9**	−0.91	**−0.50**	−0.70	**−0.28**
Lys76	−1.8	**0**	−3.0	**−1.2**	−2.6	**−0.8**	−2.01	**−0.17**	−3.43	**−1.58**
Met53	−3.0	**0**	−3.4	**−0.3**	−3.6	**−0.5**	−3.09	**−0.07**	−3.32	**−0.30**
Ser295	−1.0	**0**	−1.3	**−0.3**	−1.3	**−0.3**	−1.45	**−0.40**	−1.04	**0.01**

Table S4. Average partial atomic charges (*e*) from Mulliken population analysis along the path and polarization by Met53 (bold). Atom names in Blue = protonated phosphite, Green = water, Red = His292, Purple = nicotinamide part of NAD^+

	R		TS1		INT		TS2		P	
P	2.695	**0.003**	2.858	**0.004**	2.889	**0.003**	2.856	**-0.001**	2.653	**0.000**
H1	-0.613	**-0.003**	-0.757	**-0.001**	-0.809	**-0.001**	-0.757	**0.000**	0.092	**0.000**
O1	-1.267	**0.001**	-1.306	**0.001**	-1.331	**0.001**	-1.171	**0.002**	-1.254	**0.001**
O2	-1.241	**-0.001**	-1.295	**-0.002**	-1.362	**-0.002**	-1.342	**-0.002**	-1.253	**0.000**
O3	-0.859	**0.001**	-0.854	**-0.001**	-0.861	**-0.001**	-0.880	**0.001**	-0.849	**0.000**
H2	0.284	**-0.001**	0.221	**-0.001**	0.208	**-0.001**	0.236	**-0.002**	0.271	**0.001**
OH2	-0.440	**0.002**	-0.731	**0.002**	-0.920	**0.002**	-0.913	**0.000**	-0.858	**-0.001**
H1T	0.222	**0.000**	0.222	**0.000**	0.188	**-0.001**	0.204	**0.000**	0.285	**-0.001**
H2T	0.218	**-0.001**	0.387	**0.000**	0.332	**0.002**	0.355	**0.000**	0.322	**-0.001**
CB	-0.153	**-0.001**	-0.155	**0.000**	-0.165	**0.000**	-0.163	**0.001**	-0.169	**0.000**
HQ1	0.065	**0.001**	0.085	**0.000**	0.088	**0.000**	0.087	**-0.001**	0.090	**0.000**
HB1	0.119	**0.001**	0.111	**-0.002**	0.125	**-0.001**	0.148	**-0.003**	0.132	**-0.001**
HB2	0.114	**0.002**	0.127	**0.001**	0.142	**0.001**	0.106	**-0.001**	0.144	**0.001**
ND1	-0.164	**0.002**	-0.160	**-0.005**	-0.139	**-0.003**	-0.155	**0.000**	-0.131	**0.000**
HD1	0.321	**0.001**	0.330	**0.000**	0.373	**0.000**	0.320	**-0.002**	0.332	**0.000**
CG	-0.091	**0.004**	-0.077	**-0.003**	-0.057	**-0.002**	-0.068	**-0.002**	-0.033	**-0.003**
CE1	-0.102	**-0.007**	-0.035	**0.000**	0.037	**0.000**	0.065	**0.002**	0.061	**-0.001**
HE1	0.206	**-0.004**	0.232	**0.005**	0.268	**0.006**	0.279	**0.000**	0.284	**0.000**
NE2	-0.275	**-0.002**	-0.238	**0.002**	-0.135	**0.000**	-0.126	**0.003**	-0.146	**-0.001**
CD2	-0.221	**0.002**	-0.176	**-0.002**	-0.107	**-0.002**	-0.083	**0.006**	-0.114	**0.002**
HD2	0.184	**0.002**	0.208	**0.000**	0.232	**0.000**	0.236	**-0.002**	0.229	**0.005**
N1N	-0.099	**0.001**	-0.096	**0.001**	-0.088	**0.000**	-0.158	**0.001**	-0.233	**0.000**
HQ2	0.390	**0.000**	0.389	**0.000**	0.378	**0.000**	0.338	**0.001**	0.302	**0.000**
C6N	0.012	**0.001**	0.047	**0.001**	0.048	**0.000**	0.105	**0.000**	-0.093	**0.000**

H6N	0.218	**0.001**	0.217	**0.000**	0.227	**0.000**	0.191	**0.001**	0.100	**0.000**
C5N	-0.168	**0.000**	-0.163	**0.000**	-0.159	**0.000**	-0.176	**0.000**	-0.224	**0.000**
H5N	0.193	**0.000**	0.191	**0.000**	0.181	**0.000**	0.164	**0.000**	0.130	**0.000**
C4N	0.022	**0.000**	0.029	**0.000**	0.036	**0.000**	0.017	**-0.001**	-0.065	**0.000**
H4N	0.185	**0.000**	0.193	**-0.001**	0.178	**-0.001**	0.165	**-0.001**	0.100	**0.000**
C3N	-0.158	**-0.001**	-0.167	**0.000**	-0.171	**0.000**	-0.194	**-0.001**	-0.235	**0.000**
C2N	0.066	**0.000**	0.047	**0.000**	0.063	**0.001**	0.053	**0.001**	-0.002	**0.000**
H2N	0.251	**0.000**	0.226	**0.000**	0.244	**0.000**	0.211	**0.001**	0.172	**0.000**
C7N	0.377	**0.001**	0.353	**0.000**	0.373	**0.001**	0.369	**0.002**	0.386	**0.000**
O7N	-0.469	**-0.002**	-0.449	**-0.001**	-0.499	**-0.001**	-0.479	**-0.007**	-0.560	**0.001**
N7N	-0.386	**-0.001**	-0.393	**0.000**	-0.383	**0.000**	-0.394	**0.002**	-0.394	**0.000**
H71N	0.257	**0.001**	0.252	**0.000**	0.257	**0.000**	0.254	**0.001**	0.266	**0.000**
H72N	0.309	**-0.001**	0.329	**0.000**	0.318	**0.000**	0.301	**0.001**	0.262	**0.000**

Table S5. Parameters (measured and derived) and Cartesian coordinates (relative to C^1) involved in the development of the imidazole(ium) model (see also Figure S4).

Parameter	**Atoms involved/ derivation**	**Value**	**Standard deviation**
Measurements			
m	$C^1 – N^{1/2}$	1.327 Å	± 0.016
N	$N^{1/2} – C^{2/3}$	1.371 Å	± 0.018
O	$C^2 – C^3$	1.345 Å	± 0.021
P	$C^{2/3} – H^{2/3}$	0.943 Å	± 0.036
Q	$N^{1/2} – X^{1/2}$	1.653 Å	± 0.327
R	$C^1 – H^1$	0.943 Å	± 0.037
α	$N^1 – C^1 – N^2$	110.775º	± 2.746
β	$C^1 – N^{1/2} – C^{2/3}$	106.759º	± 2.262
γ	$N^{1/2} – C^{2/3} – H^{2/3}$	125.440º	± 2.366
ε	$C^1 – N^{1/2} – X^{1/2}$	126.287º	± 2.431
$r_{vdWaals(X)}$	X^1 and X^2	1.645 Å	± 0.202
Derivations			
α'	0.5(180-α)	34.6125º	
α^2	180-90-α’	55.3875º	
β’	180-β-α^2	17.8535º	
β^2	180-90-β’	72.1465º	
γ’	γ-β^2	53.2935º	
ε’	ε-α^2	70.8995º	
m^x	cos(α’)m	1.0921 Å	
m^y	sin(α’)m	0.7538 Å	
n^x	sin(β’)n	0.4203 Å	
n^y	cos(β’)n	1.3050 Å	
p^x	cos(γ’)p	0.5636 Å	

p^y	sin(*γ'*)p	0.7560 Å
q^x	sin(*ε'*)q	1.5620 Å
q^y	cos(*ε'*)q	0.5109 Å

Cartesian coordinates

Atom	***x*-value (Å)**	***y*-value (Å)**	***z*-value (Å)**
C^1	0.000	0.000	0.000
C^2	-(o/2) = -0.672	$-(m^y+n^y)$ = -2.059	``
C^3	o/2 = 0.672	$-(m^y+n^y)$ = -2.059	``
N^1	$-m^x$ = -1.092	m^x = -0.754	``
N^2	$-m^y$ = 1.092	$-m^y$ = -0.754	``
H^1	0.000	r = 0.943	``
H^2	$-(m^x+p^x)$ = -1.236	$-(m^y+n^y+p^y)$ = -2.815	``
H^3	m^x+p^x = 1.236	$-(m^y+n^y+p^y)$ = -2.815	``
X^1	$-(m^x+q^x)$ = -2.654	$-m^y+q^y$ = -0.213	``
X^2	m^x+q^x = 2.654	$-m^y+q^y$ = -0.213	``
Centr.	0.000	-1.125	``

Table S6. Preparation of the PTDH model from the crystal structure 4E5K[4]: missing residues (blue), sequence / structure mismatches (black) and sidechain flips as predicted by WHATIF[20] (green). Residues for which side chain electron density was partially missing are shown in orange. Note that for the QM/MM calculations the model is truncated to a 25 Å sphere centered on the NAD^+ cofactor in chain A, so the final model includes parts of chains A, D and F. The QM/MM model was prepared from a structure earlier in the refinement process than the deposited version 4E5K which is why there are some variations from the final structure. However, none of the missing or mutated residues feature in the final QM/MM model.

Chain A	Chain B	Chain C	Chain D	Chain E	Chain F
Missing: 1; 334-336	Missing: 331-336	Missing: 329-336	Missing: 1; 332-336	Missing: 1; 323-336	Missing: 1; 328-336
K130 → Q	R37 → A	K130 → Q	K130 → Q	T35 → A	K130 → Q
H22	K130 → Q	H9	Y258 → F	R36 → G	Q182 → E
N28	K177 → A	Q135	A331 → G	W92 → A	E221 → A
N80	H12	H174	H22	L93 → A	H9
N145	N28	N211	Q29	K130 → Q	H12
H174	Q47	N219	Q47	A315 → V	H22
N211	N80	Q254	Q130	E319 → A	N28
H216	Q165	N286	Q135	H9	Q47
	H174	Q314	N145	H22	N80
	N211		Q165	N28	H174
	N323		H174	N80	N211
			N211	H118	H284
			N323	HI74	H292
				Q182	N311
					Q314
					N323

Figure S1: Sequence alignment showing the conserved Met53 and mass spectra illustrating its replacement by Nle and Sec. (A) Partial sequence alignment of several active PTDH orthologs from different organisms. The orthologs exhibit high sequence diversity, with 39-72% overall pairwise identity[3]. The PTDH from *P. stutzeri* corresponds to the wild-type PTDH from which 17X-PTDH was developed. Conserved residues are highlighted in yellow. Highlighted in red is the Met53. (B) ESI-MS spectrum of Nle-17X-PTDH digested by trypsin in an SDS-PAGE gel indicating the incorporation of 3 norleucines at position X in the peptide fragment 45-DAQAXXAFXPDR-56 (M+H = 1330) (red). Also shown are the expected m/z values of fragments containing 1, 2 or 3 methionines (M+H = 1348, 1366 or 1384). Shown as a control is the peptide fragment from the tryptic digest of 17X-PTDH, which contains 3 methionines (black). The crude spectra were deconvoluted to obtain M+H charge states. (C) MALDI-MS spectrum of SeMet-17X-PTDH digested with GluC in solution. Along with other ions, diagnostic ions were observed that correspond to the peptide fragment 39-ILRRCRDAQAXXAFXPDRVDADFLQACPE-67 containing 3 selenomethionines at position X (M+H =3513 Da). Because selenium has five stable naturally occurring isotopes, the peak is greatly broadened and has a complex isotope composition. Another ion containing Se is observed at 3495 (M+H – 18) that is a dehydration product, likely formed in the mass spectrometer. No ions are visible corresponding to 1 or 2 incorporated methionines (M+H = 3417, 3465 Da respectively).

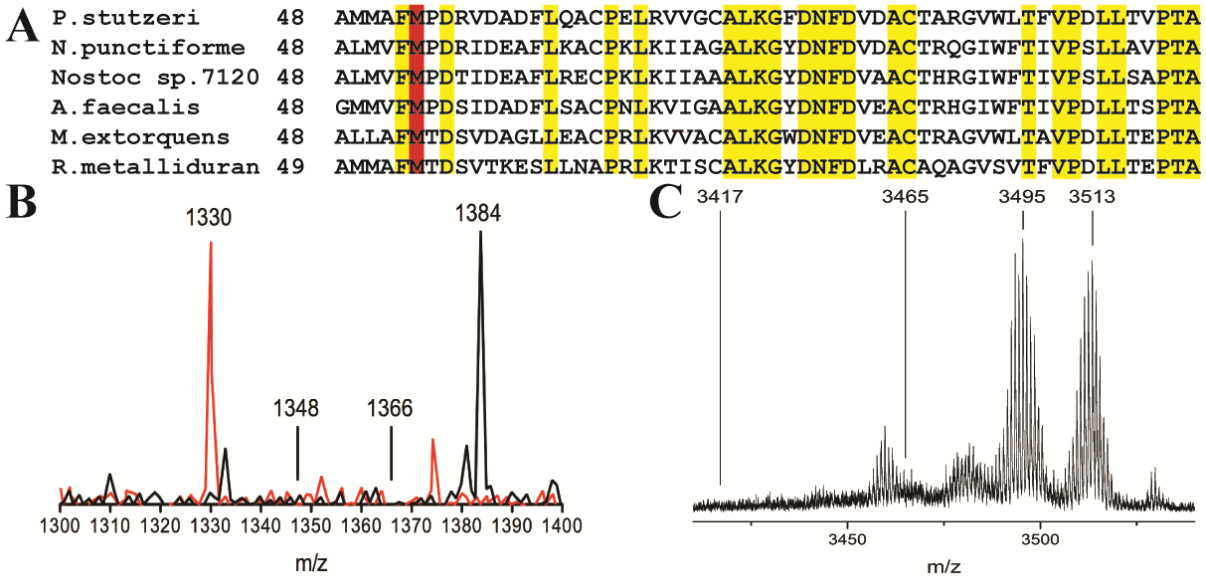


Figure S2: QM region showing atom designations.

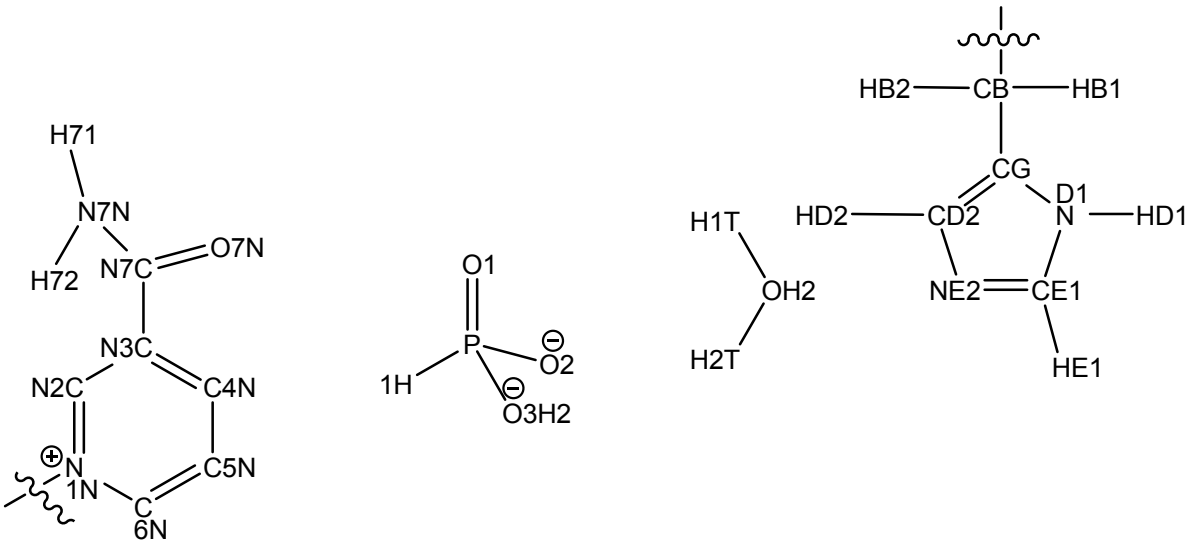


Figure S3. Perspective views of intermolecular interactions involving imidazolium rings found by an initial CSD search. (a-c): a distorted dimethylsulfoxide (DMSO) molecule in close proximity to the imidazolium ring as found in CUPNEH (a), CUPNIL (b) and CUPNOR (c). The distances shown are the shortest found between the DMSO and imidazolium entities; (d-e): interactions as found in POMBEY (d) and YESHAF (e). The distances shown are the shortest found between the imidazolium entity and the interacting atom. Hydrogen atoms and irrelevant structural components (anions, solvent molecules) have been omitted for clarity. C = grey; N = light blue; O = red; S = orange; Cl = green; I = purple.

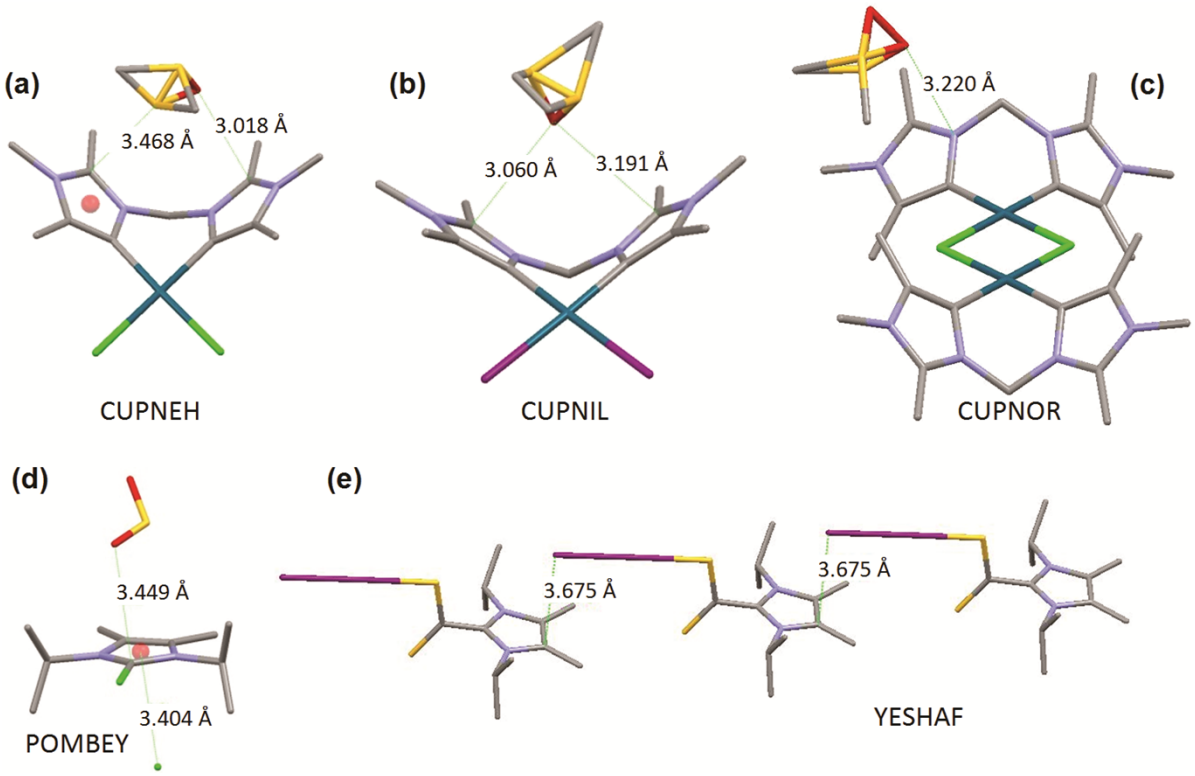


Figure S4. Model used for the statistical evaluation of imidazole/imidazolium – electron rich atom (El.R.) supramolecular interactions. (a): Flat representation of the imidazole/imidazolium model used with relevant metric parameters; (b-c): Schematic drawings of the method used to compute *x*- and *y*-values of the electron rich atom (El.R.) relative to C^1; (d) and (e) show respectively a schematic representation and a 3D space-filling model based on CSD measurements; (f) parameters used for statistical analysis: The C/N ring centroid is represented by a red sphere; α = C1 – centroid – El.R. angle; D = centroid – El.R. distance; d = ring plane – El.R. distance; r = parallel displacement parameter derived from D and d using Pythagoras' theorem. El.R. = halogen, N, O, P, S, As, Se or Te; (g-j): Illustrations of the method used to analyze the statistical preference of hits around the periphery of the imidazole/imidazolium model within a 10 Å wide and 4 Å high spherical segments centered at the *xy*-plane viewed along the *z*-axis (g) the *x*-axis (h) and perspective views (i, j); (k-m): Illustrations of the method used to analyze the statistical preference of hits around the imidazole/imidazolium model ring within a 10 Å wide and 4 Å high spherical segments centered at the *yz*-plane viewed along the *z*-axis (k) the *x*-axis (l) and perspective view (m). The red sphere illustrates the location of the ring C/N centroid defined as [0,0,0].

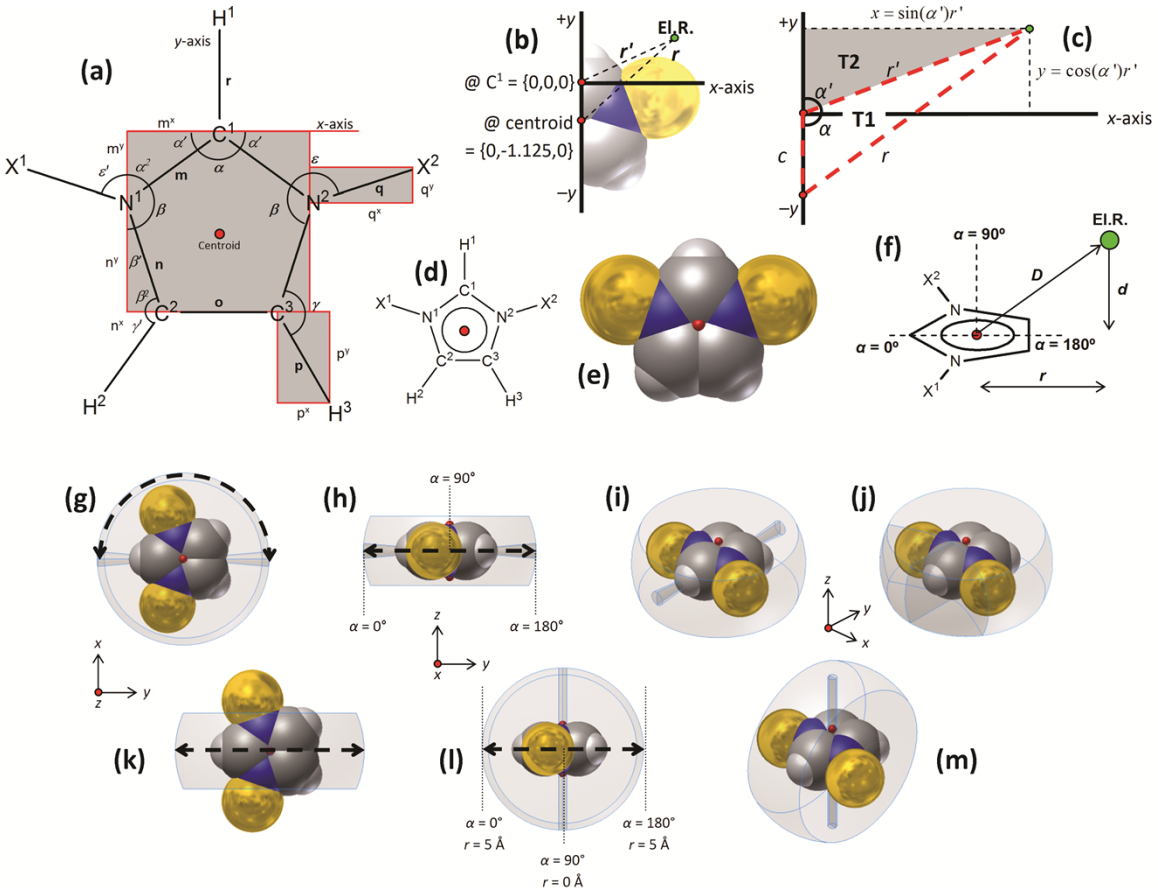
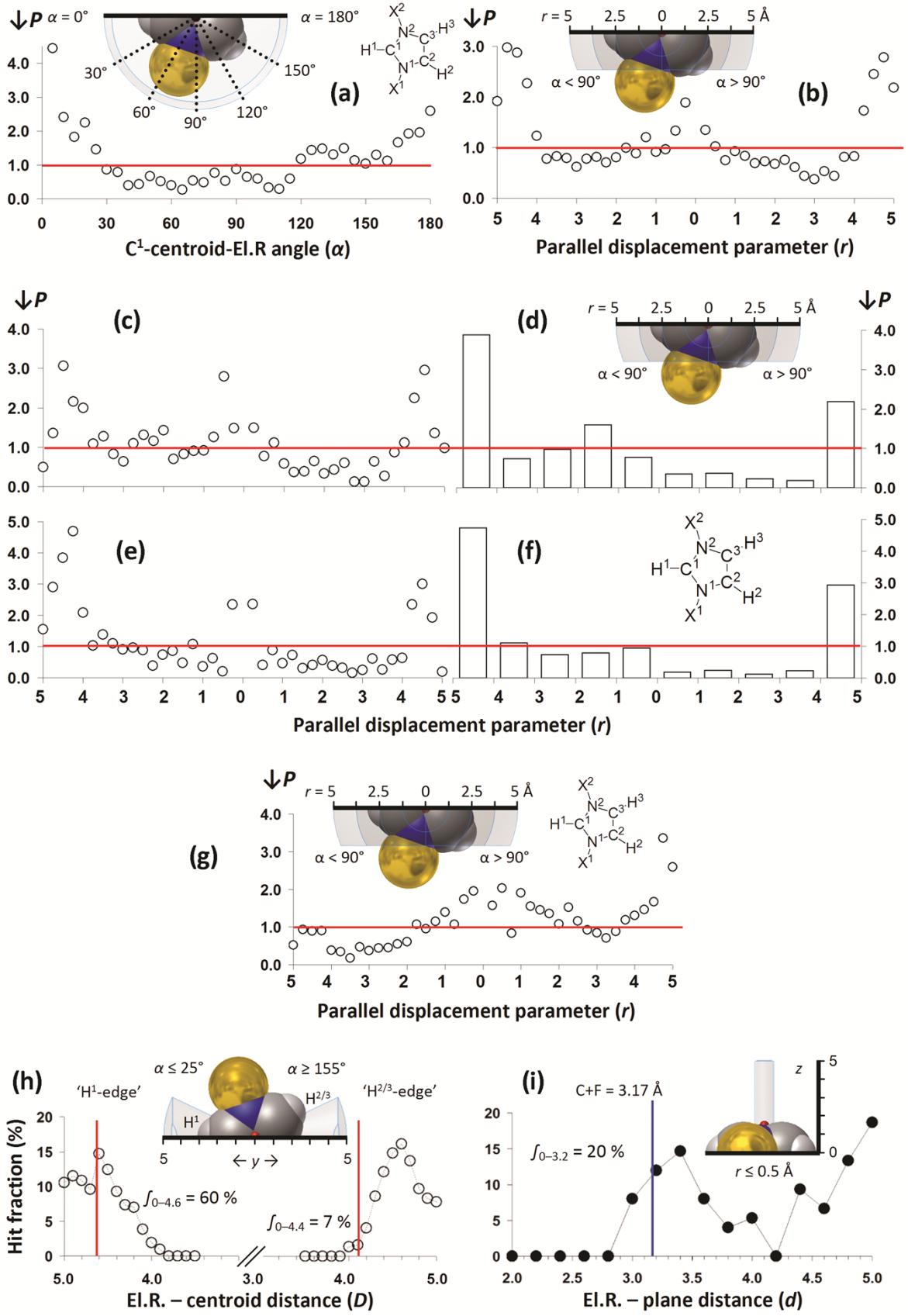


Figure S5. Results of statistical analyses: (a) Probability (*P*) as a function of *α* (see Figure S4) within a 10 Å wide and 4 Å high spherical section centered on the *xy* plane (*N* = 1985); (b) Probability (P) as a function of *r* (see Figure S4 within a 10 Å wide and 4 Å high spherical section centered on the *yz* plane (*N* = 2529); (c-g): Plot of the probability (*P*) as a function of *r* (see Figure S4) within a 10 Å wide and 4 Å high spherical section centered on the *yz* plane wherein the electron rich atom is: F, *N* = 459 (c); O, *N* = 580 (d); S, P, As, Se, Br, Te, I, *N* = 262 (e); Cl, *N* = 357 (f); N, *N* = 871 (g). (h, i): Plots of the hit fraction (in percentages) as a function of: (h) *D*, as found within the volumes characterized by |*d*| (or |*z*|) ≤ 2 Å and *α* ≤ 25º (left, N_{tot} = 365) or *α* ≥ 155º (right, N_{tot} = 313); (i) *d*, as found within the volumes characterized by *D* ≤ 5 Å and *r* ≤ 0.5 Å (N_{tot} = 75). The inset figures are shown as a guide to the eye.


↓P
α = 0°
α = 180°
30°
150°
60°
120°
90°
(a)
C¹-centroid-El.R angle (α)
r = 5 2.5 0 2.5 5 Å
α < 90°
α > 90°
(b)
Parallel displacement parameter (r)
(c)
(d)
(e)
(f)
(g)
(h)
'H¹-edge'
'H²/³-edge'
α ≤ 25°
α ≥ 155°
H¹
H²/³
← y →
∫0–4.6 = 60 %
∫0–4.4 = 7 %
Hit fraction (%)
El.R. – centroid distance (D)
(i)
C+F = 3.17 Å
∫0–3.2 = 20 %
z
r ≤ 0.5 Å
El.R. – plane distance (d)

References

1 Johannes, T. W., Woodyer, R. & Zhao, H. Directed Evolution of a Thermostable Phosphite Dehydrogenase for NAD(P)H Regeneration. *Applied and Environmental Microbiology* **71**, 5728-5734, (2005).

2 Fogle, E. J. & van der Donk, W. A. Pre-steady-state studies of phosphite dehydrogenase-demonstrate that hydride transfer is fully rate limiting. *Biochemistry-US* **46**, 13101-13108, (2007).

3 Hung, J. E. *et al.* Investigation of the Role of Arg301 Identified in the X-ray Structure of Phosphite Dehydrogenase. *Biochemistry-US* **51**, 4254-4262, (2012).

4 Zou, Y. *et al.* Crystal structures of phosphite dehydrogenase provide insights into nicotinamide cofactor regeneration. *Biochemistry-US*, 4263-4270, (2012).

5 Bas, D. C., Rogers, D. M. & Jensen, J. H. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* **73**, 765-783, (2008).

6 Hadfield, A. T. & Mulholland, A. J. Active-Site Dynamics of ASADH-A Bacterial Biosynthetic Enzyme. *Int. J. Quant. Chem. (Biophys. Q.)* **73**, 137-146, (1999).

7 Lyne, P. D., Mulholland, A. J. & Richards, W. G. Insights into Chorismate Mutase Catalysis from a Combined QM/MM Simulation of the Enzyme Reaction. *J. Am. Chem. Soc.* **117**, 11345-11350, (1995).

8 Mulholland, A. J. & Richards, W. G. Acetyl-CoA Enolization in Citrate Synthase: A Quantum Mechanical/Molecular Mechanical (QM/MM) Study. *Proteins: Struct., Funct. and Genet.* **27**, 9-25, (1997).

9 Ranaghan, K. E. *et al.* Insights into enzyme catalysis from QM/MM modelling: transition state stabilization in chorismate mutase. *Mol. Phys.* **101**, 2695-2714, (2003).

10 Ranaghan, K. E., Masgrau, L., Scrutton, N. S., Sutcliffe, M. J. & Mulholland, A. J. Analysis of classical and quantum paths for deprotonation of methylamine by methylamine dehydrogenase. *ChemPhysChem* **8**, 1816-1835, (2007).

11 Horowitz, S., Yesselman, J. D., Al-Hashimi, H. M. & Trievel, R. C. Direct Evidence for Methyl Group Coordination by Carbon-Oxygen Hydrogen Bonds in the Lysine Methyltransferase SET7/9. *J. Biol. Chem.* **286**, 18658-18663, (2011).

12 Foster, J. P. & Weinhold, F. Natural Hybrid Orbitals. *J. Am. Chem. Soc.* **102**, 7211-7218, (1980).

13 Barker, J. A. & Thornton, J. M. An algorithm for constraint-based structural template matching: application to 3D templates with statistical analysis. *Bioinformatics* **19**, 1644-1649, (2003).

14 Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242, (2000).

15 Wang, G. L. & Dunbrack, R. L. PISCES: a protein sequence culling server. *Bioinformatics* **19**, 1589-1591, (2003).

16 Heckenroth, M. *et al.* On the Electronic Impact of Abnormal C4-Bonding in N-Heterocyclic Carbene Complexes. *Chemistry-a European Journal* **15**, 9375-9386, (2009).

17 Kuhn, N., Bohnen, H., Kreutzberg, J., Blaser, D. & Boese, R. 1,3,4,5-Tetramethyl-2-methyleneimidazoline - an ylidic olefin. *J. Chem. Soc. - Chem. Comm.*, 1136-1137, (1993).

18 Kuhn, N., Bohnen, H. & Henkel, G. On the reaction of carbene adducts of carbon-disulfide with bromine and iodine. *Z. Naturforsch. B* **49**, 1473-1480, (1994).

19 http://www.ccdc.cam.ac.uk/products/csd/radii/table.php4 (July 2012).

20 Vriend, G. WHATIF. *J. Mol. Graph.* **8**, 52-56, (1990).